

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
3 May 2001 (03.05.2001)

PCT

(10) International Publication Number
WO 01/30992 A2

- (51) International Patent Classification⁷: C12N 15/00
- (21) International Application Number: PCT/US00/29139
- (22) International Filing Date: 20 October 2000 (20.10.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/161,092 22 October 1999 (22.10.1999) US
60/227,951 25 August 2000 (25.08.2000) US
- (71) Applicant (for all designated States except US): UNIVERSITY OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION [US/US]; 200 Gardner Steel Conference Center, Pittsburgh, PA 15260 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): KOIKE, Chihiro [US/US]; 5628 Hempstead Street, Pittsburgh, PA 15206-1520 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
— Without international search report and to be republished upon receipt of that report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/30992 A2

(54) Title: α 1-3 GALACTOSYLTRANSFERASE GENE AND PROMOTER

(57) Abstract: The present invention provides a recombinant expression cassette comprising an α 1-3 galactosyltransferase promoter operably linked to a polynucleotide for expression. The invention also provides a recombinant mutating cassette comprising a region of homology to an α 1-3 galactosyltransferase genomic sequence. The cassettes can be employed to express foreign genes or to disrupt the native α 1-3 galactosyltransferase genomic sequence, particularly within an animal. Thus, the invention also provides transgenic animals and methods for their production and use.

α 1-3 GALACTOSYLTRANSFERASE GENE AND PROMOTER

TECHNICAL FIELD OF THE INVENTION

This invention relates to the α 1-3 galactosyltransferase gene, promoters
5 therefor, and the use thereof to create transgenic animals.

BACKGROUND OF THE INVENTION

The current shortage of acceptable organs for transplantation is a major health concern. Because the demand for acceptable organs exceeds the supply,
10 many people die each year while waiting for organs to become available. To help meet this demand, research has been focused on developing alternatives to allogenic transplantation. Thus, for example, dialysis has been available to patients suffering from kidney failure, artificial heart models have been tested, and other mechanical systems have been developed to assist or replace failing organs.
15 Such approaches, however, are quite expensive, and the need for frequent and periodic access to such machines greatly limits the freedom and quality of life of patients undergoing such therapy.

Xenograft transplantation represents a potentially attractive alternative to artificial organs for human transplantation. The potential pool of nonhuman
20 organs is virtually limitless, and a successful xenograft transplantation would not render the patient virtually tethered to machines as is the case with artificial organ technology. Host rejection of such cross-species tissue, however, remains a major concern in this area. Some noted xenotransplants of organs from apes or old-world monkeys (e.g., baboons) into humans have been tolerated for months
25 without rejection. However, such attempts have ultimately failed due to a number of immunological factors. Even with heavy immunosuppression to suppress hyperacute rejection, a low-grade innate immune response, attributable in part to failure of complement regulatory proteins (CRPs) within the graft tissue to control activation of heterologous complement on graft endothelium, ultimately leads to
30 destruction of the transplanted organs (see e.g., Starzl, *Immunol. Rev.*, 141, 213-44 (1994)). In an effort to develop a pool of acceptable organs for xenotransplantation into humans, researchers have engineered animals producing human CRPs, an approach which has been demonstrated to delay, but not eliminate, xenograft destruction in primates (McCurry et al., *Nat. Med.*, 1, 423-27
35 (1995); Bach et al., *Immunol. Today*, 17, 379-84 (1996)).

In addition to complement-mediated attack, human rejection of discordant xenografts appears to be mediated by a common antigen: the galactose- α (1,3)-

galactose (gal- α -gal) terminal residue of many glycoproteins and glycolipids (Galili et al., *Proc. Nat. Acad. Sci. (USA)*, 84, 1369-73 (1987); Cooper et al., *Immunol. Rev.*, 141, 31-58 (1994); Galili et al., *Springer Sem. Immunopathol.*, 15, 155-171 (1993); Sandrin et al., *Transplant Rev.*, 8, 134 (1994)). This antigen is
5 chemically related to the human A, B, and O blood antigens, and it is present on many parasites and infectious agents, such as bacteria and viruses. Most mammalian tissue also contains this antigen, with the notable exception of old world monkeys and apes (including humans) (see Joziassse et al., *J. Biol. Chem.*, 264, 14290-97 (1989) and references cited therein)). The antigen is highly
10 immunogenic in humans, and many individuals show significant levels of circulating IgG with specificity for gal- α -gal carbohydrate determinants (see, e.g., Galili et al., *J. Exp. Med.*, 162, 573-82 (1985), Galili et al., *Proc. Nat. Acad. Sci. (USA)*, 84, 1369-73 (1987)). Thus, in hopes of better understanding barriers to xenotransplantation, recent attention has turned to the enzyme mediating the
15 formation of gal- α -gal moieties: α 1-3 galactosyltransferase.

The expression of α 1-3 galactosyltransferase is regulated both developmentally and in a tissue-specific manner. The cDNA for this enzyme has been isolated from many species, including pigs (Hoopes et al., poster presentation at the 1997 Xenotransplantation Conference, Nantes France; Katayama et al., *J. Glycoconj.*, 15(6), 583-99 (1998); Sandrin et al., *Xenotransplantation*, 1, 81-88
20 (1994), Strahan et al., *Immunogenics*, 41, 101-05 (1995)), mice (Joziassse et al., *J. Biol. Chem.*, 267, 5534-41 (1992)), and cows (Joziassse et al., *J. Biol. Chem.*, 264, 14290-97 (1989)). While authors have proposed to eliminate the gene from xenograft donor animals (Sandrin et al. (1994), *supra*; U.S. Patent 5,821,117
25 (Sandrin et al.)), gene knock-out procedures generally require knowledge of the genomic structure and sequence beyond the cDNA of a given gene. The genomic organization of the mouse α 1-3 galactosyltransferase homologue has been deduced (Joziassse et al., *J. Biol. Chem.*, 267, 5534-41 (1992)), and human homologues are known to be inactive pseudogenes (see Joziassse et al., *J. Biol. Chem.*, 266, 6991-98 (1991); Larsen et al., *J. Biol. Chem.*, 265, 7055-61 (1990)).
30 However, the genomic organization of an α 1-3 galactosyltransferase homologue from a species that could serve as a xenograft donor for human recipients has yet to be deduced, and no promoter for any α 1-3 galactosyltransferase homologue gene is known. As such, there exists a need for methods and reagents for
35 facilitating xenotransplantation between species, particularly between species exhibiting differential expression of the gal- α -gal epitope.

BRIEF SUMMARY OF THE INVENTION

The present invention provides a recombinant expression cassette comprising an α 1-3 galactosyltransferase promoter operably linked to a polynucleotide for expression. The invention also provides a recombinant
5 mutating cassette comprising a region of homology to an α 1-3 galactosyltransferase genomic sequence. The cassettes can be employed to express foreign genes or to disrupt the native α 1-3 galactosyltransferase genomic sequence, particularly within an animal. Thus, the invention also provides transgenic animals and methods for their production and use. These aspects of the
10 invention, as well as additional inventive features, will be apparent from the accompanying drawing, sequence listing, and the following detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A through 1I depict the genomic organization porcine α 1-3
15 galactosyltransferase gene. Figure 1A depicts all introns and exons of the gene, indicating the size of the respective elements. Figures 1B through 1I depict alternatively spliced variants isolated from pig aortic endothelial cells.

Figure 2 depicts the organization of a portion of the porcine α 1-3 galactosyltransferase promoter.

20 Figure 3 depicts the organization of the alternate splicing patterns observed in the expression of the human untranslated α 1,3 galactosyltransferase pseudogene.

DETAILED DESCRIPTION OF THE INVENTION

25 In a first aspect, the present invention provides a recombinant expression cassette in which an α 1-3 galactosyltransferase promoter is operably linked to a polynucleotide for expression. The expression cassette is "recombinant" in that within the inventive cassette, the polynucleotide for expression is other than one encoding α 1-3 galactosyltransferase. The promoter and the polynucleotide are
30 "operably linked" in that an event at the promoter (e.g., binding of cellular transcription factors and other DNA binding proteins) precipitates expression (i.e., transcription) of the polynucleotide. So long as this operable linkage is maintained, the cassette can include elements other than the α 1-3 galactosyltransferase promoter and the polynucleotide for expression. For
35 example, the cassette can contain polyadenylation sequences, repressors, enhancers, splice signals, signals for secretion (see, e.g., U.S. Patent 4,845,046 and European Patent EP-B-319,641), etc. Moreover, the expression cassette can

include more than one polynucleotide operably linked to the α 1-3 galactosyltransferase promoter, (e.g., multiple coding sequences separated by internal ribosome entry sites).

- 5 The α 1-3 galactosyltransferase promoter can be derived from any species normally expressing the gene. Thus, for example, the promoter can be derived from the bovine, porcine, or murine α 1-3 galactosyltransferase genes. Examples of such promoters are set forth at SEQ ID Nos:1-6. However, the α 1-3 galactosyltransferase promoter is not limited to one of these sequences, as it can be an active fragment of one of these sequences or a derivative of one of these
- 10 sequences having one or more mutations (e.g., point mutations, substitutions, insertions, deletions, etc.). Furthermore, given the instant disclosure, it is within the ordinary skill of the art to assay regions of the α 1-3 galactosyltransferase gene unrelated to SEQ ID NOs:1-6 for promoter activity, and the inventive expression cassette can include any α 1-3 galactosyltransferase promoters so identified.
- 15 Suitable promoters can be readily identified by construction an expression cassette in which the derivative sequence is operably linked to a desired reporter gene (e.g., RNA for detection by Northern hybridization, or DNA encoding CAT, luciferase, green-fluorescent peptide, β -galactosidase, etc.) and introducing the cassette into a suitable environment for transcription and (where appropriate) translation.
- 20 Subsequently, promoter activity is detected by assaying for the presence of the reporter by standards methods (e.g., Northern hybridization, Southern hybridization, enzymatic detection, immunohistochemistry, etc.).

- Within the expression cassette, the α 1-3 galactosyltransferase promoter can be operably linked to any desired coding polynucleotide. Generally, where
- 25 expression of a given gene or factor is desired, the skilled artisan will be in possession of the sequence of the coding polynucleotide. Thus, the polynucleotide can be expressed as a bioactive RNA molecule (e.g., an antisense RNA or a ribozyme). Alternatively, the polynucleotide can encode a protein of interest, and in this embodiment, the polynucleotide can be or comprise cDNA or genomic
- 30 DNA.

- Where the polynucleotide encodes a protein, any desired protein can be so encoded, and it need not be syngenic to the species from which the promoter is derived. Thus, for example, the cassette can be employed in animals to produce proteins facilitating growth or bulking of the animal (e.g., bovine or human growth
- 35 factor) for conferring resistance to disease or parasites. Other encoded proteins can be enzymes such as sulfo- or glycosyltransferases, (e.g., a fucosyltransferase, a galactosidase, a galactosyltransferase, a, a β -acetylgalactosaminyltransferase, an

N-acetylglycosaminyltransferase, an N-acetylglucosaminyltransferase, a sialyltransferase, etc.). Where the expression cassette is employed to generate tissue or organs for xenotransplantation into an organism lacking gal- α -gal antigens (as described below), preferably the polynucleotide encodes a Type I
5 fucosyltransferase, a Type II fucosyltransferase, an α 2-3 sialyltransferase, or an α 2-6 sialyltransferase from any species, the coding sequences of which are known (see, e.g., Larsen et al., *Proc. Nat. Acad. Sci. (USA)*, 87, 6674-78 (1990); Kelly et al., *J. Biol. Chem.*, 270(9), 4640-49 (1995), *J. Biol. Chem.*, 268(30), 22782-87 (1993), Weinstein et al., *J. Biol. Chem.*, 262(36), 17735-43 (1987)).

10 The expression cassette can be constructed by conventional methods of molecular biology (e.g., direct cloning by ligation, site specific recombination using recombinases, such as the flp recombinase or the cre-lox recombinase system (reviewed in Kilby et al. *Trends Genet.*, 9, 413-21 (1993)), homologous recombination, and other suitable methods). Typically, the promoter sequence is
15 introduced into a vector 5' (i.e., "upstream") of the coding polynucleotide and any other elements (e.g., ribosome entry sites, polyadenylation sequences, etc.), after which the construct is subcloned and grown in a suitable host organism (e.g., yeast, bacteria, etc.) from which it can be isolated or substantially (and typically completely) purified by standard methods. Thus, the invention provides a vector
20 (preferably an isolated or substantially purified vector) including a recombinant expression cassette as set forth above. Such a vector can be any desired type of vector, such as naked DNA vectors (e.g., oligonucleotides or plasmids); viral vectors (e.g., adeno-associated viral vectors (Berns et al., *Ann. N.Y. Acad. Sci.*, 772, 95-104 (1995)), adenoviral vectors (Bain et al., *Gene Therapy*, 1, S68
25 (1994)), bacteriophages, baculovirus vectors (see, e.g., Luckow et al., *Bio/Technology*, 6, 47 (1988)), herpesvirus vectors (Fink et al., *Ann. Rev. Neurosci.*, 19, 265-87 (1996)), packaged amplicons (Federoff et al., *Proc. Nat. Acad. Sci. USA*, 89, 1636-40 (1992)), papilloma virus vectors, picornavirus
30 vectors, polyoma virus vectors, retroviral vectors, SV40 viral vectors, vaccinia virus vectors) or other vectors (e.g., a cosmid, a yeast artificial chromosome (YAC), etc.). Of course, the vector can (and typically does) contain elements in addition to the expression cassette that are appropriate to the type of vector (e.g., origins of replication, marker genes, genes conferring resistance to antibiotics, etc.). The insertion of the expression cassette can disrupt one or more of these
35 elements, if desired, or the cassette can be inserted between genetic elements to minimize perturbation of the backbone vector.

Where the vector is a viral vector, preferably it is replication incompetent. Thus, for example, an adenoviral vector preferably has an inactivating mutation in at least the E1A region, and more preferably in region E1 (i.e., E1A and/or E1B) in combination with inactivating mutations in region E2 (i.e., E2A, E2B, or both
5 E2A and E2B), and/or E4 (see, e.g., International Patent Application WO 95/34671). An AAV vector can be deficient in AAV genes encoding proteins associated with DNA or RNA synthesis or processing or steps of viral replication (e.g., capsid formation) (see U.S. Patents 4,797,368, 5,354,768, 5,474,935, 5,436,146, and 5,681,731). Where the vector is a retroviral vector, the cis-acting
10 encapsidation sequence (E) essential for virus production in helper cells can be deleted upon reverse transcription in the host cell to prevent subsequent spread of the virus (see, e.g., U.S. Patent 5,714,353). Where the vector is a herpesvirus, inactivation of the ICP4 locus and/or the ICP27 cassette renders the virus replication incompetent in any cell not complementing the proteins (see, e.g., U.S.
15 Patent 5,658,724, see also DeLuca et al., *J. Virol.*, 56, 558-70 (1985); Samaniego et al., *J. Virol.*, 69(9), 5705-15 (1996)).

To use the inventive recombinant expression cassette, it is introduced into a eukaryotic cell in a manner suitable for the cell to express the coding polynucleotide. A vector harboring the recombinant expression cassette is
20 introduced into a eukaryotic cell by any method appropriate for the vector employed, which generally are well-known in the art. Thus, plasmids are transferred by methods such as calcium phosphate precipitation, electroporation, liposome-mediated transfection, microinjection, viral capsid-mediated transfer, polybrene-mediated transfer, protoplast fusion, etc. Viral vectors are best
25 transferred into the cells by infecting them.

Depending on the type of vector, it can exist within the cell as a stable extrachromosomal element (which can even be heritable, see e.g., Gassmann, M. et al., *Proc. Natl. Acad. Sci. (USA)*, 92, 1292 (1995)) or it can integrate into the host cell's chromosomes. Thus, the invention provides a chromosome including a
30 recombinant expression cassette such as described above, as well as a cell including such a cassette (and such a chromosome). The $\alpha 1$ -3 galactosyltransferase promoter of the expression cassette can be native to such a cell or chromosome, or it can be exogenous to the cell or chromosome. Where the promoter is native to the cell or chromosome, preferably the polynucleotide for
35 expression within the cassette (the non-native polynucleotide) displaces the operable linkage between the native polynucleotide encoding $\alpha 1$ -3 galactosyltransferase such that it is no longer operably linked to the native $\alpha 1$ -3

galactosyltransferase promoter. Such displacement can be accomplished where the non-native polynucleotide is cloned between the promoter and the native polynucleotide (i.e., upstream of the native polynucleotide), especially where the non-native polynucleotide contains one or more transcriptional termination signals (preferably in all three putative reading frames). Of course, the non-native polynucleotide also can be introduced into the locus such that it destroys the native exon/intron boundaries and/or introduces inactivating mutations (e.g., deletions, insertions, frame-shifts, etc.) into the native coding sequence.

Preferably, the transgenic cell presents a suitable microenvironment for the coding polynucleotide within the expression cassette to be expressed. In many instances, the transgenic cells can be used to study the tissue specificity, dynamics, and kinetics of the promoter, for example by assaying for the expression of the polynucleotide within the cells. However, as the absence of activity is as useful as the presence of promoter activity in these contexts, any cell can be employed for such purposes; such a cell can be *in vivo* or *in vitro*. Preferably, the cell is derived from a species syngenic to the source of the promoter so that, by virtue of the properties of the α 1-3 galactosyltransferase promoter present within the expression cassette, the polynucleotide within the cassette is expressed within such transgenic tissues, organs, or animals with the same kinetics and tissue specificity as the native α 1-3 galactosyltransferase gene in wild-type animals. Where the cells are *in vivo*, they are typically cells of a mammal (e.g., human cells), and can be any type of cells. Suitable cells for use *in vitro* include yeast, protozoa (e.g., *T. cruzi* epimastigotes), cells derived from any mammalian species (e.g., VERO, CV-1, COS-1, COS-7, CHO-K1, 3T3, NIH/3T3, HeLa, C1271, BS-C-1 MRC-5, etc.), insect cells (e.g., *Drosophila* Snyder cells), or other such cells. In other applications, the cell can be employed to construct transgenic tissues, organs, or animals, as described below, in which case the cell typically is a spermatozoon, ovum, zygote, primordial germ cells, or embryonic stem cell.

In another embodiment, the invention provides a method of mutating a region of a chromosome comprising an α 1-3 galactosyltransferase gene. In accordance with the inventive method, a recombinant mutating cassette comprising a region of homology to the α 1-3 galactosyltransferase gene is recombined with a chromosome which has an α 1-3 galactosyltransferase gene such that homologous recombination occurs between the cassette and the chromosome. As a result of the homologous recombination, a mutation is introduced into the native α 1-3 galactosyltransferase chromosomal gene sequence.

Thus, the final step of the method involves screening for successful recombination.

The inventive method employs a recombinant mutating cassette including at least a first region of homology to an α 1-3 galactosyltransferase genomic sequence, and the invention provides such a cassette. Within such a cassette, this first region of homology is adjacent to either to at least one polynucleotide for insertion or to a second region of homology. The mutating cassette is "recombinant" in that neither the second region of homology nor the polynucleotide for insertion is adjacent to the first α 1-3 galactosyltransferase genomic sequence in its native state (i.e., within a chromosome).

The insertion cassette can include more than one polynucleotide for insertion and/or more than one region of homology to all or a portion of the α 1-3 galactosyltransferase genomic sequence. Indeed, where the cassette includes a region for insertion, preferably it has at least two regions of homology flanking the region for insertion. Where more than one region of homology is present, whether adjacent to each other or flanking a region for insertion, the cassette can be used to replace any span of the target chromosomal genomic sequence that lies between the two homologous chromosomal regions. Where multiple regions of homology are present, they should generally be arrayed in the same 5' to 3' orientation relative to one another.

A region of homology can be homologous to any portion of the genomic sequence of an α 1-3 galactosyltransferase gene or the antisense strand thereof. The region can be homologous to the gene of any desired species, such as those discussed above, and it can be homologous to an intron, an exon, a promoter sequence, or any other desired sequence from the genomic DNA. To this end, regions of homology can be selected from the promoter sequences disclosed in SEQ ID NOs:1-6. Alternatively (or additionally) a region of homology can be selected from a portion of the genomic sequence from an α 1-3 galactosyltransferase homologue. In this light, some of the murine sequences have been published (see, e.g., Joziassse et al., *J. Biol. Chem.*, 267, 5534-41 (1992)), and additional portions are set forth as SEQ ID NOs: 17-25. Portions of the porcine genomic sequence are disclosed herein as SEQ ID NOs: 7-16. Portions of the human α 1,3 galactosyltransferase pseudogene genomic sequences are set forth at SEQ ID NOs: 35-42, and various (untranslated) human cDNA transcripts are set forth as SEQ ID NOs: 27-34, and those from Rhesus monkeys are set forth at SEQ ID NOs: 43-44. These sequences disclosed herein, as well as the published murine sequences, include the intron/exon boundaries from which one of skill in

the art can isolate additional intronic genomic sequences by techniques such as genome walking, 5' RACE, 3' RACE, etc.

A region of homology to the genomic sequence of an $\alpha 1$ -3 galactosyltransferase gene need not be an exact complement to the genomic sequence; however, the region must be sufficiently homologous to the $\alpha 1$ -3 galactosyltransferase gene to permit homologous recombination between the cassette and the genomic DNA *in vivo*. Indeed, in some embodiments (e.g., for introducing point mutations into the genomic sequence), a region of homology preferably contains some mismatched bases. Thus, typically, the region of homology will bear at least about 75 % homology to a portion of the $\alpha 1$ -3 galactosyltransferase gene or its antisense strand (such as at least about 85 % homology to a portion of the $\alpha 1$ -3 galactosyltransferase gene or its antisense strand), and more typically the region of homology will bear at least about 90 % homology to a portion of the $\alpha 1$ -3 galactosyltransferase gene or its antisense strand (such as at least about 95 % or even at least about 97 % homology to a portion of the $\alpha 1$ -3 galactosyltransferase gene or its antisense strand). Any commonly employed method (e.g., BLAST database searching) for calculating percent homology can be used to select a suitable region of homology. Similarly, while the length of the region of homology is not critical, it should be sufficiently long to facilitate homologous recombination between the cassette and the genomic DNA *in vivo*. Thus, typically the region of homology will be at least about 50 nucleotides long (such as at least about 75 or 100 bases long), and more typically it will be at least several hundred bases long (such as at least about 250, 500, or even 750 bases long). Indeed, in many applications, the region of homology preferably is several thousand bases long to maximize the likelihood of homologous recombination *in vivo*. The ideal length of a region of homology depends in part on the number of such regions within the cassette – where one or few regions of homology are present, they should be longer to facilitate recombination between the cassette and the genomic DNA; conversely, where the cassette contains several regions of homology, they can be shorter without reducing the likelihood of recombination events.

Where present within the cassette, a region for insertion can be or comprise any DNA which is desired to be introduced into the genomic sequence of an $\alpha 1$ -3 galactosyltransferase gene. Thus, the region can comprise genetic regulatory elements (e.g., enhancers, promoters, repressors, etc., the sequences of which are known) or consensus binding sites for DNA-binding proteins (e.g., restriction endonucleases, transcription factors, etc.). In many applications, a region for

insertion can comprise a polynucleotide for expression, such as those set forth above, or even expression cassettes. A preferred polynucleotide for insertion is an expression cassette for expressing a positive marker flanked by FRT sites, thus facilitating the identification of chromosomes into which the polynucleotide for insertion has integrated as well as excision of the cassette.

The mutating cassette can be constructed by any desirable molecular techniques, and typically, the mutating cassette will be engineered within a vector, such as those set forth above. Typically, the vector is a gene transfer vector suitable for introducing the cassette into a host cell. In addition to the region(s) of homology and the polynucleotide for insertion elements, the mutating cassette can have other components, such as, for example, an expression cassette, a region of homology to other genes or chromosomal regions, a polyadenylation sequence, etc., and it is preferred that the insertion cassette comprises a cassette for expressing at least one marker gene (which may be or comprise the polynucleotide for insertion). Such a marker can be either positive (conferring a visible phenotype to the cells) or negative (killing cells or rendering non-recombinant cells growth-impaired), and both can be used in conjunction. Examples of such positive and negative selection markers are the neosporin resistance (neo^R) gene, the hydromycin resistance (hyg^R) gene, and a thymidine kinase gene (e.g., HSV tk); other suitable markers are known in the art (see, e.g., Mansour et al., *Nature*, 336, 348-52 (1988); McCarrick et al., *Transgen. Res.*, 2, 183-90 (1993)). A marker gene sequence can be bordered at both ends by FRT DNA elements, and/or with stop codons for each of the three putative reading frames being inserted 3' to the desired DNA sequence. Presence of the FRT elements permits the marker to be deleted from the targeted chromosome, and the stop codons ensure that the $\alpha 1,3$ galactosyltransferase gene remains inactivated following deletion of the selectable marker, if inactivation is the desired result of the use of the mutating cassette. The relative orientations of the positive and negative selectable markers are not critical. However, where a positive marker is employed, it should be located between regions of homology, while any negative marker should be outside the regions of homology, either 5' or 3' to those regions.

In accordance with the inventive method, homologous recombination occurs between the $\alpha 1-3$ galactosyltransferase genomic chromosomal DNA and the region (or regions) of homology in the mutating cassette. Where more than one region of homology is present in the cassette, any portion of the genome lying between the homologous target sequences is replaced by whatever sequence lies between the regions of homology in the cassette. Thus, where the mutating

cassette contains a region for insertion flanked by two regions of homology, it will be introduced into the genomic sequence adjacent to the sites of homology, replacing that portion of the genomic sequence. Of course, where the two flanking regions of homology are normally adjacent to each other in the chromosomal sequence, the region for insertion is introduced into the chromosome without replacing any native sequence. Similarly, where no region for insertion is present within the cassette, that portion of the chromosome lying between the two regions of homology in the cassette is deleted as a result of the recombination events. Where the cassette contains a region of homology that differs slightly from the homologous sequence within the genome, it can be employed to introduce point mutations into the genomic sequence.

While the recombination event can occur *in vitro*, typically such homologous recombination occurs within a host cell between an exogenous vector containing the cassette and a chromosome within the host cell containing an α 1-3 galactosyltransferase genomic sequence. Thus, the present invention provides a cell harboring a mutating cassette, as described above. The vector can be introduced into the host cell by any appropriate method, such as set forth above. Commonly, however, the vector is introduced into small cells (e.g., embryonic stem cells) by electroporation and into large cells (e.g., ova or zygotes) by microinjection. Where microinjection is employed, the vector preferably is injected directly into a nucleus or pronucleus of the cell.

The last step in the method is to screen for successful recombination events. Any assay to detect such events can be employed in the context of the inventive method. In accordance with one such assay, chromosomal DNA is screened by PCR or Southern hybridization. For example, where the mutating cassette is designed to delete a portion of the α 1-3 galactosyltransferase genomic sequence, the absence of signal using a probe or primer directed against the region to be deleted indicates a positive recombination event. Conversely, where the cassette includes a region for insertion, a positive result using a probe or primer directed against the region for insertion is indicative of a positive recombination event. Of course, the chromosomal DNA can be sequenced to confirm the correct insertion/deletion/replacement. Where recombination is directed within cells, the events can be screened by assaying for any markers present in the mutating cassette.

By employing the inventive method, one of skill in the art can use the inventive mutating cassette to introduce targeted deletions, insertions, or replacement mutations into any predefined site within the α 1-3

galactosyltransferase genomic sequence. Any desired amount or portion of the gene can be thus deleted, which can lead to complete inactivation of the gene. For introducing inactivating mutations into the gene, preferably at least one region of homology is selected to recombine with the promoter (to inactivate it) or exons 4-9, which contain the coding sequences. Similarly, the inventive method can introduce functional expression cassettes in place of the α 1-3 galactosyltransferase gene, which can be under the control of the native α 1-3 galactosyltransferase promoter or an exogenous promoter within the cassette (especially where the native α 1-3 galactosyltransferase promoter is destroyed). Thus, the present invention provides a recombinant chromosome containing such a mutation, and a recombinant cell comprising such a chromosome.

As mentioned above, the invention provides recombinant cells and chromosomes comprising a recombinant expression cassette comprising an α 1-3 galactosyltransferase promoter or a mutating cassette, as described above. Indeed, as a result of using these reagents and methods, the invention also provides a cell having a mutant α 1-3 galactosyltransferase genomic sequence, as described above. While any cell having such exogenous genetic sequences is within the scope of the invention, preferably the cells are suitable for constructing a recombinant animal, and are most preferably totipotent cells. Thus, preferred cells are embryonic stem (ES) cells, ova, primordial germ cells (PGCs), and zygotes. ES cells and PGCs are especially preferred because such cells can be obtained and cultured in relatively large numbers relative to ova and zygotes. Using such cells, a transgenic animal having an expression cassette comprising an α 1-3 galactosyltransferase promoter or a disruption in this gene can be constructed by methods known in the art (see e.g., U.S. Patents 5,850,004 (MacMicking et al.), 5,942,435 (Wheeler), 5,523,226 (Wheeler), and 5,175,383; White et al., *Transplant. Int.*, 5, 648-50 (1992); McCurry et al., *Nat. Med.*, 1, 423-427 (1995); Hogan et al., *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1986); Hammer et al., *Nature*, 315, 680 (1985); Murray et al., *Reprod. Fert. Devl.*, 1, 147, (1989); Purselet et al., *Vet. Immunol. Histopath.*, 17, 303 (1987); Rexroad et al., *J. Reprod. Fert.*, 41, (suppl.), 119 (1990); Rexroad et al., *Molec. Reprod. Devl.*, 1, 164 (1989); Simon et al., *BioTechnology*, 6, 179 (1988); Vize et al., *J. Cell. Sci.*, 90, 295 (1988); Wagner, J. *Cell. Biochem.*, 13B (suppl.), 164 (1989); Thomas et al., *Cell*, 51, 503 (1987); Capecchi, *Science*, 244, 1288 (1989); Joyner et al., *Nature*, 338, 153 (1989); Ausubel et al., *Cur. Prot. Mol. Biol.*, John Wiley & Sons (1987)).

Where ova and zygotes are employed, after the introduction of the cassette, they can be implanted into surrogate mothers to develop into adult animals.

Where ES cells or PGCs are employed, after the introduction of the cassette, they typically are further manipulated (e.g., by injection into a blastocyst or morula, co-culture with a zona pellucida-disrupted morula, fusion with an enucleated zygote, etc.) such that their mitotic descendants are found in a developing embryo. Such an embryo typically is a chimera composed of normal embryonic cells as well as mitotic descendants of the introduced ES cells or PGCs. Alternatively, the genome of an ES cell or PGC can be incorporated into an embryo by fusing the ES cell/PGC with an enucleated zygote to create a non-chimeric embryo in which all nuclei are mitotic descendants of the fused ES cell/PGC nucleus. In any event, to produce a transgenic animal, the embryo or zygote is implanted into a pseudopregnant animal, which, after suitable gestation, gives birth to an animal containing the mutant chromosome containing the cassette in its germ line (if a chimera) or possibly all of its cells. Of course, as mentioned above, where the animal is engineered to include a non-mutating expression cassette, it can be inherited as an extrachromosomal plasmid (Gassmann, M. et al., *supra*)). However constructed, the presence of the recombinant allele can be confirmed by performing Northern hybridization or rt-PCR on RNA isolated from the animal in question.

After birth and sexual maturation, a chimeric animal can be mated to generate a heterozygous animal comprising a disrupted α 1-3 galactosyltransferase gene or recombinant expression cassette (integrated or extrachromosomal) including a α 1-3 galactosyltransferase promoter. Heterozygotes can be crossed to produce a homozygous strain. Such animals having a recombinant expression cassette including an α 1-3 galactosyltransferase promoter, as discussed above, will express the polynucleotide for expression of such cassette within the same tissue types and with the same kinetics as a wild-type animal of the same species and strain expresses the α 1-3 galactosyltransferase gene. Of course, homozygous transgenic animals of the present invention having a disruption in the α 1-3 galactosyltransferase gene will produce altered forms of the protein or no functional protein at all. Desirably, the phenotype of such "knock out" animals relative to an animal having a wild type α 1-3 galactosyltransferase gene is a markedly increased time of survival of cells isolated or derived from the transgenic animal in the presence of human serum, which can be assessed by any desired method (see, e.g., Osman et al., *Proc. Nat. Acad. Sci. (USA)*, 94, 14677-82 (1997)).

The inventive transgenic animals are useful for any use to which animals can be put, and they can be any desired species (e.g., pigs, cows, mice, cats, dogs, etc.). Transgenic mice in which a reporter gene is operably linked to the α 1-3 galactosyltransferase promoter are valuable reagents for assessing the activity and specificity of the promoter. Transgenic livestock (e.g., pigs, cows, goats, and the like) having an inventive expression cassette in which a growth hormone is expressed under the control of the α 1-3 galactosyltransferase promoter can be matured or bulked better than commonly employed strains. Tissue obtained from a transgenic animal according to the present invention can be implanted into a host according to standard surgical methods, and the invention concerns a method of xenotransplantation from a transgenic animal as described herein. The invention also provides a transgenic organ consisting essentially of transgenic cells engineered as described above (e.g., a lung, a heart, a liver, a pancreas, a stomach, an intestine, a kidney, a cornea, skin, etc.), particularly for use in the method of transplantation. The host can be any animal host, such as a pig, a dog, a cat, a cow, a goat, etc. Of course, the recipient can be a human as well, in which case the source animal preferably is a pig.

Transgenic animals lacking a functional α 1-3 galactosyltransferase gene are attractive sources of organs and tissues for xenotransplantation into primates, especially humans, because the tissues of such animals lack the highly antigenic gal- α -gal epitope. Similarly, transgenic pigs having a recombinant expression cassette in which a coding sequence for Type I fucosyltransferase, a Type II fucosyltransferase (especially α (1,2) fucosyltransferase), an α 2-3 sialyltransferase, or an α 2-6 sialyltransferase is operably linked to the α 1-3 galactosyltransferase promoter also are suitable sources of xenotransplantation tissues, as these encoded enzymes compete for the same substrate as α 1-3 galactosyltransferase, and their presence can reduce (preferably below an antigenic threshold) the gal- α -gal antigens in tissues derived from such animals. Indeed, α (1,2) fucosyltransferase converts this substrate into the universally-tolerated H antigen (i.e., the "O" blood-type antigen) and also blocks the addition of the α 1,3 gal moiety. As such, a gene encoding α (1,2) fucosyltransferase is an especially preferred polynucleotide for expression to be included within the inventive recombinant expression cassette. A preferred source animal for xenotransplantation tissues (and by extension the tissues themselves) preferably contains a disruption in the α 1-3 galactosyltransferase gene as well as having a recombinant expression cassette in which a coding sequence for Type I fucosyltransferase, a Type II fucosyltransferase (especially α (1,2)

fucosyltransferase), an α 2-3 sialyltransferase, or an α 2-6 sialyltransferase is operably linked to the α 1-3 galactosyltransferase promoter. More preferably, the animal contains a disruption in the native promoter of α 1-3 galactosyltransferase and an α (1,2) fucosyltransferase coding sequence under the control of its own
5 promoter. Most preferably, the source animal also expresses exogenous human complement regulatory proteins, as discussed above, to further minimize host resistance of the xenograft tissue.

It will be apparent that a transgenic animal created in accordance with the invention can have the exogenous gene cloned in place of the native α 1,3
10 galactosyltransferase gene (i.e., a "knock-in" approach). Indeed, in many embedment such a "knock-in" approach is preferable, for example to avoid the potential of the development of congenital cataracts in purely "knock-out" animals (e.g., as a result of opportunistic infections of microbes bearing the gal- α -gal motif). Indeed, such an approach can afford a safe alternative to broadband
15 antibiotics in livestock and pets, a current public health concern. In this respect, the invention can be employed to create heartier and healthier livestock and pets.

While one of skill in the art is fully able to practice the instant invention upon reading the foregoing detailed descriptions, in conjunction with the drawing and the sequence listing, the following examples will help elucidate some of its
20 features. In particular, these examples indicate how the genomic structure of the porcine α 1-3 galactosyltransferase gene is elucidated, and how the identity and activity of the α 1-3 Galactosyltransferase promoter is assessed. As these examples are presented for purely illustrative purposes, they should not be used to construe the scope of the invention in a limited manner, but rather should be seen
25 as expanding upon the foregoing description of the invention as a whole.

Many experiments described in these examples employed well known techniques and reagents (see, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2d edition, Cold Spring Harbor Press (1989)). Accordingly, in the interest of brevity, the examples do not present the experimental protocols in
30 detail. In the experiments, enzymatic isolation and culture of porcine aortic endothelial cells (PAEC) was performed. PAEC were maintained in Dulbecco's modified essential medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 10,000 units of Heparin (ELKINS-SINN, Inc., Cherry Hill, NJ), 15 mg of endothelium growth supplement (Collaborative Biomedical Product Inc., Bedford,
35 MA), L-glutamine, and penicillin-streptomycin. RNA was obtained from the organs of pigs (Brain, Heart, Spleen, Gut, and Thymus) and PAEC using Trizol reagent

(Gibco Ltd.). Primers used to clone and identify regions of the porcine, murine, human, and Rhesus monkey genes are set forth at SEQ ID NOs: 45-96.

Example 1

5 This example describes the identification of the 5' untranslated region and genomic structure of the porcine α 1-3 galactosyltransferase gene.

A comparison of published sequences for the α 1-3 galactosyltransferase cDNA (Hoopes et al., *supra*, Katayama et al., *supra*; Sandrin et al., *supra*; and Strahan et al., *supra*) revealed a divergence in the 5' boundary. Some of these
10 cDNA contain putative 5' untranslated sequences that bear a high (> 70 %) homology to murine sequences identified as the second exon, and it was hypothesized that this region is conserved as an exon in the porcine genome as well.

Further 5' sequence was cloned using 5' RACE, and the putative
15 transcription initiation site was probed by S1 protection assay, using standard protocols. Briefly, a plasmid containing the upstream genomic sequence was digested with restriction enzyme, Pml I, and linearized. The DNA was phosphorylated with shrimp alkaline phosphatase, heated to inactivate the enzyme, and then precipitated with ethanol. The linearized plasmid was digested again
20 with Bgl II to yield a probe fragment, which was then end-labeled with γ -³²P-ATP.

The probe was purified using G-25 sephadex, and about 16 μ l was mixed with 20 μ g of total RNA from pig aortic endothelial cells (PAEC), pig brain, and yeast (control), and the aliquots were coprecipitated using NH₄OAc and ethanol. Pellets were resuspended in a standard hybridization buffer, heated to 95 °C for 3-
25 4 minutes, and then incubated at 42 °C overnight.

After incubation, the yeast sample was split into two aliquots, and to each was added a standard S1 nuclease buffer. S1 nuclease was added to one aliquot, while the other did not receive the enzyme. The PAEC and brain samples each received the enzyme and the buffer. All samples were incubated for 30 minutes at
30 37 °C, after which the reactions were stopped by the addition of a standard S1 inactivation buffer. Following the reaction, the samples were then precipitated, resuspended in 5 μ l of a standard gel loading buffer, and resolved using a 6% denaturing polyacrylamide gel.

The data revealed at least 8 separate alternatively spliced transcripts from
35 PAEC, and additional splicing patterns from brain transcripts. Analysis of these sequences revealed three potential upstream exons (1, 1A, and 2), the boundaries of which comply with the AG-GT consensus, and six coding exons (4-9) also were

identified, which agreed with published results. Interestingly, the pig sequence seemingly lacks upstream exon 3 of the mouse 5' untranslated region. The overall organization of the pig genome is depicted in Figure 1. Alternatively spliced forms isolated from PAEC are indicated in Figures 1B through 1I. Exon 1A is
5 observed in transcripts isolated from brain tissue.

As mentioned, the transcripts obtained from PAEC and brain revealed several alternative splicing patterns. Using the genomic clone, intronic sequences were identified by "gene walking" using the method and reagents supplied with the UNIVERSAL GENEOMEWALKER™ KIT (Clontech Labs., Inc.). Primers
10 (Seq ID NOs:41-56) were designed to hybridize with the cDNA, and also to the adapter sequence supplied with the Clontech kit. A series of nested PCR reactions was then performed to clone SEQ ID NOs:7-16, which were sequenced. From these results, the intron/exon boundaries were elucidated.

Summing the nucleotides of all identified exons predicts a transcript of
15 about 3.8 kb. This prediction was assessed by Northern analysis. 20 µg of total RNA from PAEC, and pig brain, heart, spleen, gut, and thymus, were respectively separated on formamide agarose gels, and electrotransferred onto nylon membrane. The blots were hybridized with radiolabeled probes (2.5-4.0 x 10⁴ cpm/ml) specific for pig GT exon 1 and exon 9 identified. The blots were exposed
20 to Bio-MAX films (Eastman Kodak Co., Rochester, NY) for 6 days with intensifying screen. The results revealed primary transcripts of between 3.5-3.8 kb, in accordance with the predicted size and the published size for the bovine transcript.

25 Example 2

This example describes the identification of the 5' untranslated region and organization of the murine α 1-3 galactosyltransferase gene.

To identify the 5' and 3' ends of α 1,3GT gene transcripts, 5'- and 3'- RACE procedures were performed using the Marathon cDNA Amplification Kit
30 (Clontech) with the spleen poly A⁺ RNA of Balb/C adult male as template. To identify exon-intron boundaries or 5'- and 3'-flanking region of the transcripts, Murine GenomeWalker libraries were constructed using the Universal GenomeWalker Library Kit (Clontech) with Balb/C genomic DNA.

The results of these experiments revealed several genomic sequences,
35 which are set forth at SEQ ID NOs: 17-25. The deduced 5' untranslated nucleotide sequences are longer by 56 bp than previously reported (Joziassse et al., *J. Biol. Chem.*, 267, 5534-41 (1992)). The relative intensity of Luciferase activity

by the pGL3/1280 construct was 15-fold higher than that of pGL3-Basic. The 3'-RACE revealed an extended 3'-UTR sequence 30 bp more than previously reported (*Id.*), but no other 3' UTR exon usage. The overall length of the transcript was 2586 bp, 89 bp longer than previously reported (*Id.*).

- 5 An overall comparison of 5'-UTR of cDNA sequences of the porcine (747 bp) and murine (492 bp) α 1,3GT gene indicates that the homology is observed only in the region of exon 2 (71.7%). Exon 3 observed in mice is not observed in the pig. Murine exon 1 shows no homology with porcine exon 1.

10 **Example 3**

This example describes the identification of the organization of the human and Rhesus monkey α 1-3 galactosyltransferase untranslated pseudogene.

- Working from published partial sequence of the human α 1,3 GT ninth exon, primers were designed to identify the start and end of the gene by 5'-RACE, 3'RACE and rtPCR, as described above. Several alternate transcripts were identified, and these are set forth as SEQ ID NOs:27-34. The sequences were compared to those of other species employing a formula based on the consensus motif of the splicing acceptor junction: total number of pyrimidines plus 1 (for a branched A) among forty nucleotides per junction. Intron exon boundaries were confirmed as discussed above (see SEQ ID NOs: 35-42). The organization of the alternative splicing patterns observed is indicated in Figure 3.

- Using similar techniques, primers were designed based on a partial published sequence (Genbank Accession No. M73306) having homology to exon 9. Initially, 3'RACE showed only poly-A tails, evidence that transcripts exist. 5'-RACE results revealed sequences of high homology to those α 1,3 sequences previously identified (e.g., porcine, bovine and murine), consistent with the identity of the sequence as the Rhesus pseudogene. The sequence of the Rhesus monkey transcripts are set forth at SEQ ID NOs: 43 and 44.

30 **Example 4**

This example describes the identification of the porcine, murine, and bovine α 1-3 galactosyltransferase promoters.

- Using PCR and restriction digestions, various sized fragments between nucleotides 1981 and 2992 of SEQ ID NO:7 (porcine) and between nucleotides 375 and 1325 (murine) were generated. The fragments were cloned into a plasmid such that they were operably linked to a luciferase coding sequence. PAEC were then transfected with these constructs and probed for luciferase activity, along

with a positive and a negative (no promoter) control. All fragments exhibited significantly greater promoter activity over the negative control (between about 15% and 90 % relative light units, as compared to the positive control, the negative control exhibiting no luciferase activity). These results indicate that the regions are promoters and that the 5'-RACE results discussed in Examples 1 and 2 most likely represent the potential transcription initiation site (TIS). Moreover, sequence analysis of these regions reveals the presence of at least 8 SP1 or GC boxes within it and potentially seven AP-2 consensus binding motifs (see also Figure 2). This suggests that the gene may contain alternative start sites, and that sequences within exon 1 may also contain promoter activity. Other sequences from which α 1,3 GT promoters can be derived are set forth as SEQ ID NOs:1-6.

All of the references cited herein, including patents, patent applications, and publications, are hereby incorporated in their entireties by reference.

While this invention has been described with an emphasis upon preferred embodiments and illustrative examples, it will be obvious to those of ordinary skill in the art that variations of the preferred embodiments may be used and that it is intended that the invention may be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications encompassed within the spirit and scope of the invention as defined by the following claims.

WHAT IS CLAIMED IS:

1. A recombinant expression cassette comprising an α 1-3 galactosyltransferase promoter operably linked to a polynucleotide for expression, other than a polynucleotide encoding α 1-3 galactosyltransferase.
- 5 2. The recombinant expression cassette of claim 1, wherein said promoter is derived from the bovine, porcine, or murine α 1-3 galactosyltransferase genes.
3. The recombinant expression cassette of claim 1 or 2, wherein said promoter comprises any of SEQ ID Nos:1-6.
4. The recombinant expression cassette of any of claims 1-3, wherein said
10 promoter comprises an active derivative of any of SEQ ID Nos:1-6.
5. The recombinant expression cassette of any of claims 1-4, wherein said polynucleotide for expression encodes an antisense RNA molecule or a ribozyme.
6. The recombinant expression cassette of any of claims 1-5, wherein said polynucleotide for expression encodes a protein.
- 15 7. The recombinant expression cassette of claim 6, wherein said protein is a fucosyltransferase, a galactosyltransferase, a β -acetylgalactosaminyltransferase, an N-acetylglycosaminyltransferase, an N-acetylglucosaminyltransferase, a sialyltransferase, or a sulfotransferase.
8. The recombinant expression cassette of claim 6, wherein said protein is a
20 Type I fucosyltransferase, a Type II fucosyltransferase, an α 2-3 sialyltransferase, or an α 2-6 sialyltransferase.
9. The recombinant expression cassette of any of claims 6-8, wherein said polynucleotide for expression is heterogenic to said promoter.
10. The recombinant expression cassette of claim 9, wherein said
25 polynucleotide for expression is human and wherein said promoter is porcine.
11. The recombinant expression cassette of any of claims 6-8, wherein said polynucleotide for expression is a cDNA.
12. The recombinant expression cassette of any of claims 6-8, wherein said polynucleotide for expression is genomic DNA.
- 30 13. A recombinant mutating cassette comprising a first region of homology to an α 1-3 galactosyltransferase genomic sequence adjacent to either a second region of homology to said α 1-3 galactosyltransferase genomic sequence or a polynucleotide for insertion.
14. The recombinant mutating cassette of claim 13, comprising first and
35 second regions of homology to an α 1-3 galactosyltransferase genomic sequence flanking a polynucleotide for insertion.

15. The recombinant mutating cassette of claim 13 or 14, wherein a region of homology is homologous to an exon, an intron, or a promoter of said α 1-3 galactosyltransferase genomic sequence.

16. The recombinant mutating cassette of any of claims 13-15, wherein a
5 region of homology is homologous to all or a portion of any one of SEQ ID NOs: 1-42.

17. The recombinant mutating cassette of any of claims 13-16, wherein said polynucleotide for insertion comprises an expression cassette.

18. The recombinant mutating cassette of claim 17, wherein said
10 expression cassette encodes a marker.

19. A vector comprising the recombinant cassette of any of claims 1-18.

20. The vector of claim 19, which is an oligonucleotide, a plasmid, a cosmid, or a virus.

21. A transgenic cell harboring the vector of claim 19 or 20.

22. A chromosome comprising the recombinant expression cassette of any
15 of claims 1-18.

23. A transgenic cell harboring the chromosome of claim 22.

24. The transgenic cell of claim 23, wherein said α 1-3 galactosyltransferase promoter is native to said cell.

25. The transgenic cell of claim 23 or 24, wherein said polynucleotide for
20 expression displaces a native polynucleotide encoding α 1-3 galactosyltransferase.

26. The transgenic cell of claim 23 or 24, wherein said polynucleotide for expression is cloned between said promoter and a native polynucleotide encoding α 1-3 galactosyltransferase.

27. The transgenic cell of claim 26, wherein said polynucleotide for
25 expression comprises a stop codon.

28. The transgenic cell of any of claims 21, or 23-27, which is an embryonic stem cell, an ovum, a primordial germ cell, a spermatozoon, or a
zygote.

29. The transgenic cell of any of claims 21, or 23-27, which expresses said
30 polynucleotide for expression.

30. The cell of claim 29, wherein said polynucleotide for expression encodes a Type I fucosyltransferase, a Type II fucosyltransferase, an α 2-3 sialyltransferase, or an α 2-6 sialyltransferase, and wherein said cell produces said
35 protein.

31. The transgenic cell of any of claims 21 or 23-30, wherein said cell produces a heterogenic complement regulatory protein (CRP).

32. The transgenic cell of claim 31, wherein said CRP is human and wherein said cell is nonhuman.

33. An embryo consisting essentially of transgenic cells according to any of claims 21 or 23-32.

5 34. An organ consisting essentially of transgenic cells according to any of claims 21 or 23-32.

35. The organ of claim 34, which is a lung, a heart, a liver, a pancreas, a stomach, an intestine, a kidney, or skin.

10 36. A transgenic animal consisting essentially of transgenic cells according to any of claims 21 or 23-32.

37. The transgenic animal of claim 30, which is a cattle, a mouse, a pig, a cat or a dog.

15 38. A transgenic knockout animal comprising a homozygous disruption in an endogenous α 1-3 galactosyltransferase gene, wherein said disruption prevents the expression of a functional α 1-3 galactosyltransferase protein.

39. The transgenic knockout animal of claim 38, wherein cells isolated from said knockout animal exhibit an increased time of survival in the presence of human serum relative to comparable cells isolated from an animal having a wild type α 1-3 galactosyltransferase gene.

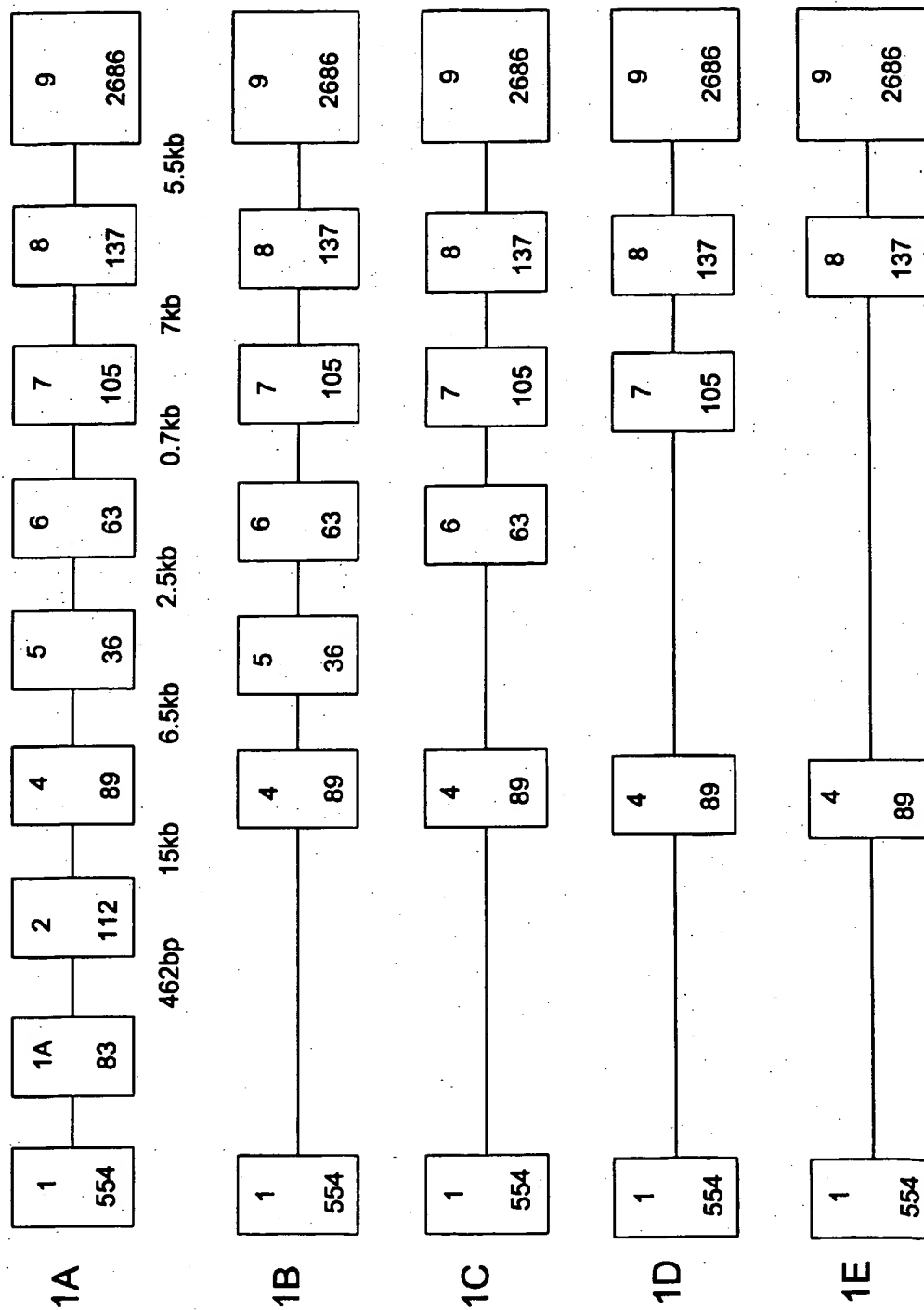
20 40. The transgenic knockout animal of claim 38 or 39, wherein the insertion replaces DNA at the start of the coding region of said α 1-3 galactosyltransferase protein.

41. The transgenic knockout animal of claim 38 or 39, wherein the insertion replaces the promoter of said wild type α 1-3 galactosyltransferase gene.

25 42. The transgenic knockout animal of any of claims 38-41, which produces at least one human protein selected from the group of proteins consisting of α 1-3 galactosyltransferase, α (1,2) fucosyltransferase, and complement regulatory proteins.

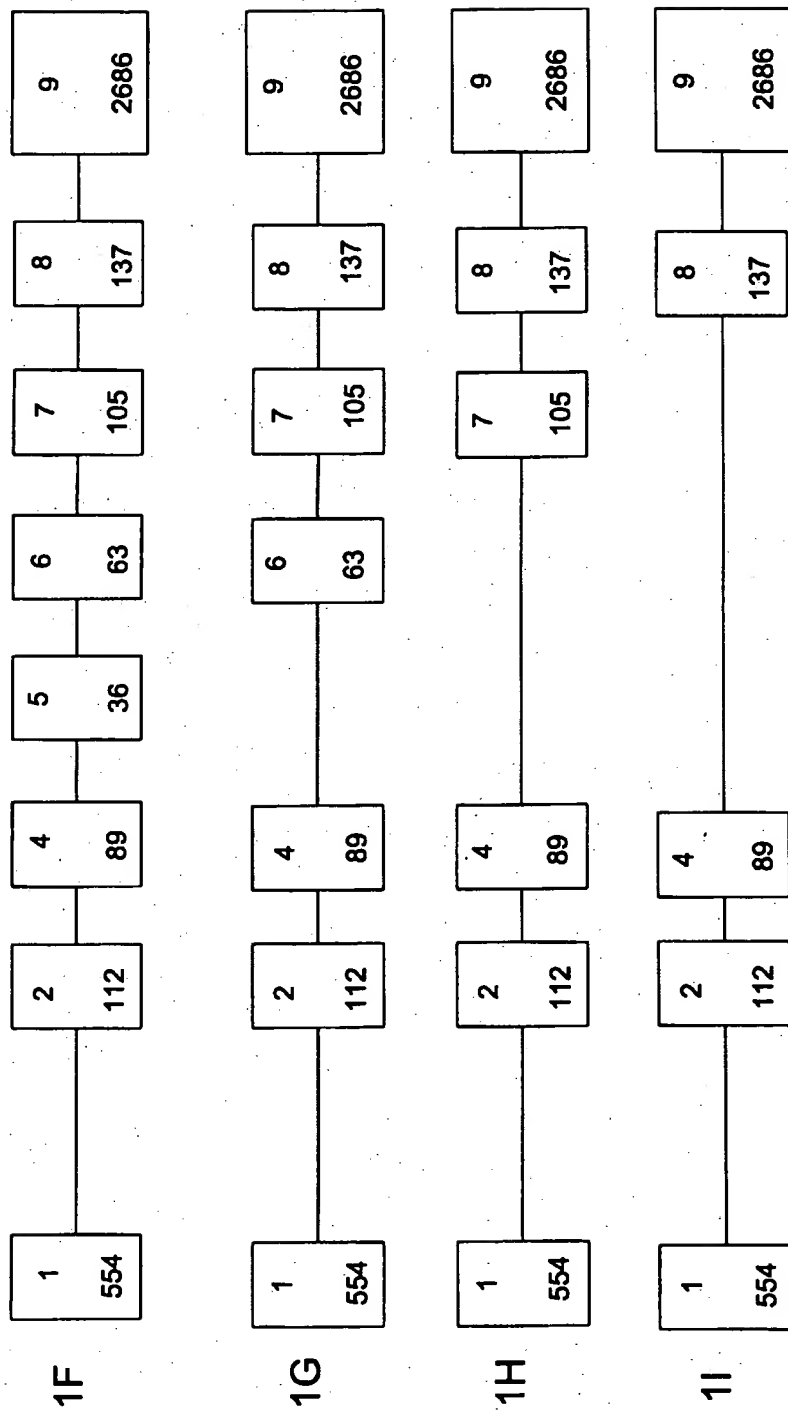
30 43. The transgenic knockout animal of any of claim 38-42, which is a pig.

1 / 4



SUBSTITUTE SHEET (RULE 26)

2 / 4



SUBSTITUTE SHEET (RULE 26)

FIG. 2

Schematic genomic organization of human $\alpha 1, 3$ GT gene

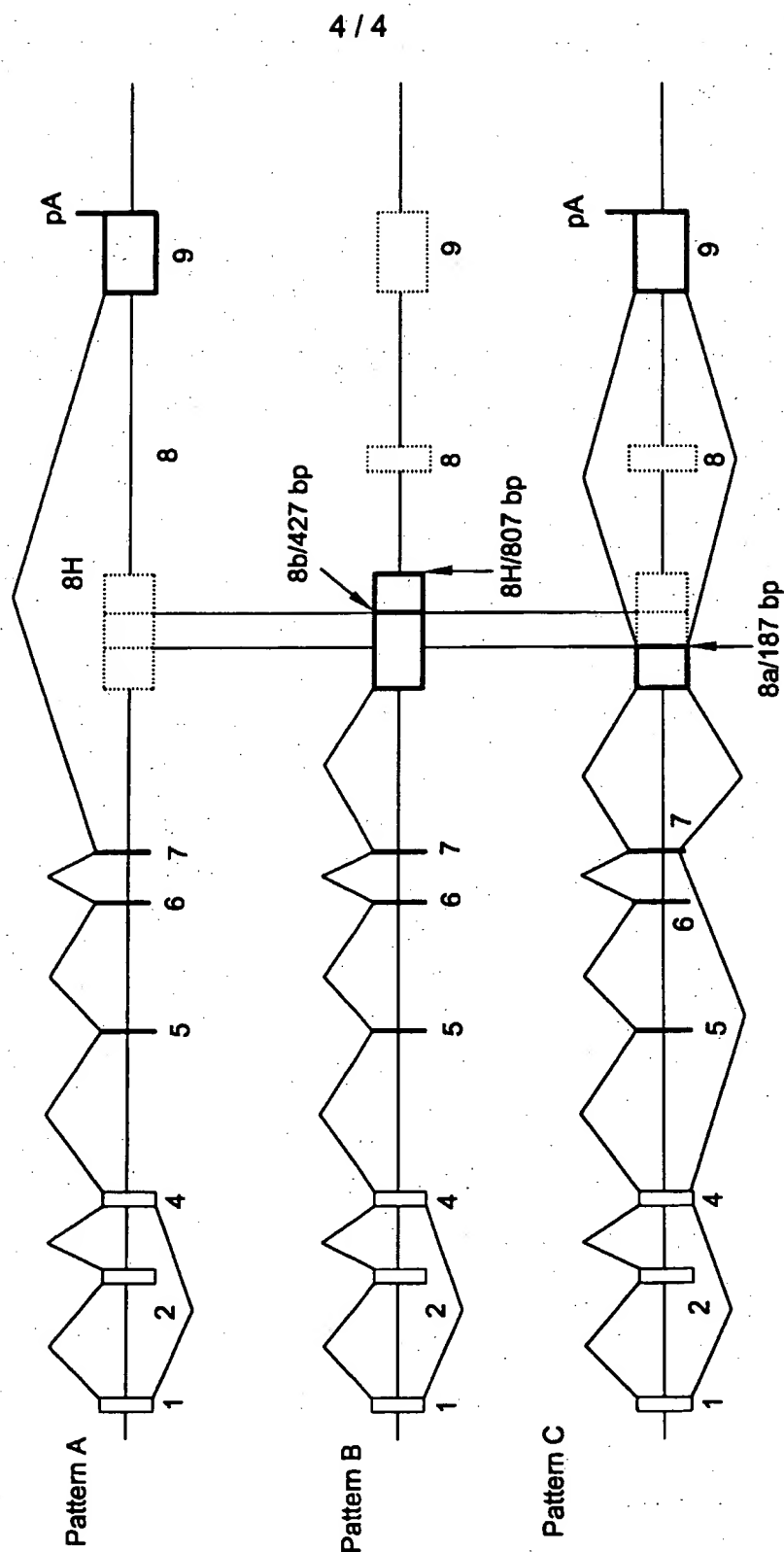


FIG. 3

SEQUENCE LISTING

<110> University of Pittsburgh of the Commonwealth System of Higher Education

Koike, Chihiro

<120> alpha 1,3-galactosyltransferase gene and promoter

<130> 206780

<150> US 60/227,951

<151> 2000-09-25

<150> US 60/161,092

<151> 1999-10-22

<160> 96

<170> PatentIn version 3.0

<210> 1

<211> 1117

<212> DNA

<213> Sus scrofa

<400> 1

agatctctgt tcttttcaaa tcaggatgaa acagttaaaa ttatacatca cactcaggtt 60

ctgtgccatt ttcatgtcac aattccaatg ccttaaaata ttaagaaac taatttctta 120

gtctctgaag tcccgtgggtg aatgatcctg gcaaaagcaa gttctgaatt ttgcagcagt 180

aaaatagatg gtccgggacc ccaaggagtc ttgtaaaggc tgagtgaggg cagccggatg 240

tgacctacacc agctcatcag aagtgaactg ttgtcacact gggcactaaa gcaccaactc 300

tgaaatataa tttttgatta tgttcctcc taaaataact aaagcacaaa ctctgaaata 360

taatttttcgt ttacgtttctc tccctctact aatattccag cagagaacag agcccgcgcc 420
aggtgtccag taccagccc ctcatatccg aagctcagga cttgggggtt tcgggagaga 480
gcggctccag cgcgtcgggt tgtagctact gcatctgtgc tttccttcc ccaggaaaca 540
aatggtggat cggacctccc aggtctctcg cgccccgcca cccctcccg tgtagcagg 600
gcgcaggggt ccggggcccc tccctgcagt actgggtgat agacccact ccaccctcg 660
gtccctcca cccccaccac gtgcaggcca gagaaggcaa agaggcccag ccaccctcac 720
caggaattt cttttctttt tttgctggtt tcaggctttt ttctgcctga gtgaaaatga 780
aacaacacc ccctgcgct cccggccacc agacacacac gcgcaccggc actcgcgcac 840
tcgcgcctc ggctcttag cggcgtgtc tggggcggga cccgctctgc acaacagcc 900
gcgggcggg tggagcggg agctcgccg ccgcccga gtgcccgcg gttcctcgc 960
gcccctgcc gccaccccg aggagcacac agcggccggc gggccggagc gcaggcggca 1020
caccccgccc cggcacgccc tgccgagctc aggagcacgc cgcgcgccac tgttcctca 1080
gccgaggac ccgcccggg gccgggagcc gaggtgt 1117

<210> 2

<211> 900

<212> DNA

<213> Sus scrofa

<400> 2

ttgtcacact gggcactaaa gcaccaactc tgaaatataa tttttgatta tgttcctcc 60
taaaataact aaagcacaaa ctctgaaata taattttcgt ttacgtttctc tccctctact 120
aatattccag cagagaacag agcccgcgcc aggtgtccag taccagccc ctcatatccg 180

aagctcagga cttgggggtt tcgggagaga gcggctccag cgcgtcgggt tgtagctact 240
gcattctgtgc tcttccttcc ccaggaaaca aatggtggat cggacctccc aggctcttcg 300
cgccccgcca cccctccccg tgtagcagg gcgcagggtt ccggggcccc tccctgcagt 360
actgggtgat agacccact ccacctccg ggtccctcca cccccaccac gtgcaggcca 420
gagaaggcaa agaggcccag ccacctcac cagggaattt cttttctttt ttgctggtt 480
tcaggctttt ttctgctga gtgaaaatga acaaacacc cctgcgct cccggccacc 540
agacacacac gcgcaccggc actcgcgcac tcgcgccctc ggctcctag cggccgtgtc 600
tggggcggga cccgctctgc acaaacagcc gcgggccggg tggagcggg agctcgccgc 660
ccgcgcacca gtgcccgcg gcttcctcgc gccctgcc gccaccccg aggagcacac 720
agcgcccggc gggccggagc gcaggcggca cccccgcc cggcacgcc tgccgagtc 780
aggagcacgc cgcgcgccac tgttcctca gccaggagc ccgccccggg gccgggagcc 840
gaggtgtggg ccattccccg gcgcaccag cttctgccga tcaggtgggt cccgctgggc 900

<210> 3

<211> 1938

<212> DNA

<213> Sus scrofa

<400> 3

gaggaagggc aacatcagac ccaatggttc ctatcagat ttgttaacca ctgagcctcg 60
atgggaactc ctgggtgctt gcttcttgaa aggaccagtt tatcttagcc cagttcctga 120
gcctccaaat gctgtgaact ttccctccca gttgaccaca gtccagctgc ctgcatcatt 180

taatgtgaaa gatcttcctt gagtccgtac ttaggtgctc tgtggtgctt ggtattggg 240
cggtgaaccc aagagaagga aaaaacgggg tctatccacg accctgtggc cctgagaccc 300
tgtagactca ggggaagtca gaattcccaa gagaaggcag cttccagcag gaagatttct 360
gtgcatcttt gtttttaaca cacacactga aagggaatgt ttgtgaggca ttttcccaag 420
gtggacacac ctgcataacc actacctggc tcgagaaaca acatgacaag cccccccccc 480
tccccagca gctctctgag cctccccttc ccagtctcta ccactccac tctgacttct 540
ggcaccacag attggttttg tctttttttt ttttttgtct ttttagggct acacttggg 600
catatggaag ttcccaggct aggggtccaa ttggagctgt ggctgttggc ctacaccaca 660
gccacagcaa catgggatcc gagccgcac tgcaacctac accacagctg gtggcaatac 720
tggatcctta acccactgag tgaggccagg gatcgaactt gcattctcgt acatactggt 780
cagatttggt tctgctgagc caccatggga actccctggt tttgtctatt tttttttttt 840
tttttgtctt ttttgccatt tcttgggccc ctcttgccgc atatggaggt tcccaggcta 900
agggccaat cggagccgta gcccagcct acgccagagc cacagcaacg tgggatccga 960
gccgagtctg caacctacac cacagctcgc ggcaacgcca gatcccttaa cccactgagc 1020
aaggccaggg accgaacccg caacctcatg gttcttagtc ggattcgtta accactgcgc 1080
cacgacggga actcccgtt ttgtctattt ttgaacgtta aataaatgca agcatccagg 1140
gctgctttga ctcagtagca tgtgtgagat ttaccctggt gatgtcagca gctgtggctg 1200
gttccttctc acggatgtgt gtgacctca cctggaccac acctgatctg gctgatgatg 1260
ggccttgggg tttttccagc ttttggccc aggtcacgctc tctgtttgaa cttaaagca 1320

cttgctttca ggtattaatc tggggcggaa tgactggaac atgaggtgtg gttgggtcag 1380
ctttagtaca tgccagcagg gaggatttca gtagtttatt aagcagatct tgaagactgt 1440
gggtcaactag ctcatgcccc acaggagggg gcggtgaatt tcttccccag aacaggagtg 1500
acaagctaaa ttaggcattcc atccgctgga agctgagggg gcagttcttg gctcctttct 1560
gtcaggtttc ggccccttct ccttagtctg gggtttctag gctctactcc caggaagtgt 1620
ctggggccac ttgggaacaa tgggtggggg ggctctgagc ccctacttac ttcatttccc 1680
tccttcagcc aaagccccct gtgtcctctg tttacatag tggggttctg agaatgactt 1740
catttttttt tttttttttt ttaaagcttt agctgttgcg acatttaca atccactgct 1800
gtgaggtctc ttccaggtag gaaattgtat tttgggagca ggaggtgggt gtggggaggg 1860
ttaagcatta ttacagccaa gagttgggtt gggcctcagt gaccttttga agttcttata 1920
gcttggtctg ccatgcag 1938

<210> 4

<211> 820

<212> DNA

<213> Mus musculus

<400> 4

actaaccagt gagtgtagaa agcaggaggt gtcttttctt actgtagtta ggacagggcg 60
ggttggctct tcttatggac aagatggaaa aggggtgcag gtaggggcaa agtgagagac 120
actcgaattt gagagacaga cagactccta acagtgaagg aaggaccaag ccaaatcaa 180
gcctgggcaa agtctcaggc actaactttg ctgtgttggg tgatgggagg taatctcgtc 240
acaacttttc aaaccacctc gttccactg caaggagaca ccatcaagtg tttgaagatg 300

gcagggaac ctctcaaca aacacacaca caaacgtttt attattttat attatatttg 360
catgcaaagt actgtgtttc attatggcat ttccatacat atgcgattgc acaaactctt 420
gaaaatcatc caagaaacag caaagcggga aataatgttg tggggggggg gcgcggagga 480
gagagaacag agactggaga gagtgtgtc ctcttctgtc cgggggccag gaagaggcta 540
ggaggcgagg gatgtcaacg ccactagctc ctccctcagg aaggaccca gggactctta 600
ttttgtagt ttgtctgtc tgggccacta tcggccccag aacagatctg actgcctctt 660
tcattcgccc ggaggtagat aggtgtgtct taggaggctg gagattctgg gtggagccct 720
agccctgctt ttcttagct ggctgacacc ttcccttgta gactcttctt ggaatgagaa 780
gtaccgattc tgctgaagac ctgcgcctct caggctctgg 820

<210> 5

<211> 930

<212> DNA

<213> Mus musculus

<400> 5

tgacactgaa gccacgcggg ggcttcagtg gggaggagggt gtgggcgagc gcgagcgccg 60
ctattccggc ccagccctac ctcggtcctt gcttttgtcc tggtcactcg atcatttctt 120
ctgtatecac ttctgaactc taggtctgtt cccaccctga acagtgtcgc tgcattctgt 180
tgcttactgg ggtctccgc caccttcctt cgctatccga atagctgata ttcagggcag 240
cacagggcag ggcagggcag ggcagggcga gtagggcaga tcagatcctg ggaccaccgg 300
tactaaccag tgagtgtaga aagcaggagg tgtcttttcc tactgtagtt aggacagggc 360

gggttggtc ttcttatgga caagatggaa aaggggtgca ggtaggggca aagtgagaga 420
cactcgaatt tgagagacag acagactcct aacagtgaag gaaggaccaa gccaaaatca 480
agcctgggca aagtctcagg cactaacttt gctgtgttgg gtgatgggag gtaatctcgt 540
cacaactttt caaaccacct cgttcccact gcaaggagac accatcaagt gtttgaagat 600
ggcaggggaa cctctcaaca aaacacacac acaaacgttt tattatttta tatttatttt 660
gcatgcaaag tactgtgttt cattatggca tttcataca tatgcgattg cacaaactct 720
tgaaaatcat ccaagaaaca gcaaagcggg aaataatgtt gtgggggggg ggcgcgagg 780
agagagaaca gagactggag agagtgtgt cctccttgct gcgggggcca ggaagaggct 840
aggagggcgg ggatgtcaac gccactagct cctccctcag gaaggacccc agggactctt 900
atttttgtag ttttgcttgt ctgggccact 930

<210> 6

<211> 501

<212> DNA

<213> bovine

<400> 6

cctccctgtc catcaccaac tcccggagct cactcagact catgtccatc gagtcgggtga 60
tgccatccag ccattctatc ctctgtcgtc gccttctcct ctgtcccca atcccgcaca 120
gcatcagagt cttttccaat gagtcaactc ttcgcatggg gtggccaaag tactggagtt 180
tcagcttttag catcatcccc tccaaagaaa tcccagcggc cgagtccggg gcgggacccg 240
ctctgcacaa acaccggggg ccgggcccag ctgggagcgt cgagcccgt gccagcgcc 300
cgccgggtcc ctgcgcccc tgcccgcgc cccggaggag cgcccggcg ccggccgacg 360

ggagcgcagc ggcacacccc gccccggcac gcccgcgagg ctcgggagga ggcagcgcgc 420
cgactgttcc ggcagccgag gacgccgccg gggagccgag gcgccggcca gccccagcg 480
cgcccagctt ctgcggatca g 501

<210> 7

<211> 3976

<212> DNA

<213> Sus scrofa

<220>

<221> misc_feature

<222> (580)..(580)

<223> "n" is a gap of from about 600 to about 800 nucleotides

<220>

<221> promoter

<222> (1863)..(2992)

<223> fragments and derivatives of this region have promoter activity.

<220>

<221> 5'UTR

<222> (2463)..(3016)

<223> Untranslated exon 1 runs from about nucleotide 2436 to about nucleotide 3016

<220>

<221> Intron

<222> (1)..(2462)

<220>

<221> Intron

<222> (3017)..(3976)

<400> 7

aggcctaaac ctagaactcc tgaccctgaa gctaaggaat ataactctga aggtgttttc	60
cagtcagtag aataacacag agtttccaca catgcgtggg tctctttcta ggttgcttat	120
tctgttccat tggccaata aaccatcctg gcgctaagtc tatactgagt tcaactgcgtt	180
tcatgggtctg tcttggtatc tgggtgaaca agagcccaac tctccccctcc ctgctttgtc	240
aagactgcct tggttatatc tggccccctc ccgctgctgt ccaaatttta agaatagctg	300
gccaaagctcc cccaaaactc tgttggcatt tgtcttgagt ttatagggtg atgcatggag	360
aattgttgcc ttcgtgatgc tgatgcttcc cagtgtcac tcgggggtct ctttccttcc	420
acctaagac ttctgcacat ggttctgctt gggtcactct tccccagcc ttcacctagt	480
gaactcctcc tctcctggt ctcagggctc cctgcacct tatttcttcc ttagagccct	540
gatcacaatg gtcctgaaat cactcattgc gtgggtcttn gtgacagata gtaggtccca	600
gtaaatatct gttaaagaa tgaaggaagt ttaggtagga aggtcttcgg gacctggagc	660
accttgcca tagttagagg gatggtgacc agaggtactt aacttgctg tgccttggt	720
ttcttctac aaaaccgga tgtgatcaga atgtgtataa gatgaagtga gtcagctag	780
gccgtgaggc aagtggagca aagcctggca agggatcaga gctacttggt tacctgccct	840
gcccttctgc tcagtgaatc ttcagtcctg cactcctgtg atgctcctgg aggtccaac	900
actctttccc cagcagtgat ccgctctga ctccacctct cctatgaact agtcacctta	960
tttctactca gcatatgaca caaatgagtc tcaggaagaa tgactcataa ggccttaaac	1020
ctagaactcc tgaccctgaa gctaaggaat ataactctga aggtgttttc cagtcagtag	1080

aattgctagt tagatttggg gagctacata gttctcaaaa gaaaacaaaa cttccggacc 1140
cgccgtgtta atttgaatta tttttatctt attgttactg aaataggtat aaacctagaa 1200
ctaagaatga agtcctcatg ctcttagctc tgcacaccta ccatgatacc aaagcaaatac 1260
ttttaagtag gtgcaattac agccacaaaa ccaataaaat ccaaattagc aacgttaaat 1320
ttatgcaact gatgacatgg tgctgaaatc aaacctcttg cattgagtct aatggtagca 1380
gagtgatgtt ttacatgtt tcattccctg tgtcatcatc ttttgatttt gatcctgatg 1440
agctatcact tcagccatgg tcagaattac cgtcataatt ttcactaaaa aaaaaaccca 1500
aaaaacacat ttattatcca atttgatggg ctgagcaatt taaacactgg atcctcaagt 1560
gcaataatga caactgggaa atactttgct aacatcactc cttgtgtatt tatttactgc 1620
atcattaaag acctagtgc aagtgttca ccgatgacaa taatggcgca gtttatgctt 1680
ttgcaaagga tccattgttc ggattgtcat ggagctctc attcctgagc taccctgtgg 1740
ggctgatgat tcaactctcc cacccttag tccactgaac ccatcaggaa agttcattat 1800
cccaagctcc aagatgtcac ttggctccct gcagcctctc tgcaaccgct aagtattcaa 1860
tcagatctct gttcttttca aatcaggatg aaacagttaa aattatacat cacactcagg 1920
ttctgtgcc a tttcatgtc acaattccaa tgccttaaaa tatttaagaa actaatttct 1980
tagtctctga agtcccgtgg tgaatgatcc tggcaaaagc aagttctgaa ttttgagca 2040
gtaaaataga tggccggga cccaaggag tcttgtaaag gctgagtgg ggcagccgga 2100
tgtgcctaca ccagctcatc agaagtgaac tgtgtcaca ctgggcacta aagcaccaac 2160
tctgaaatat aatttttgat tatgttcct cctaaaataa ctaaagcaca aactctgaaa 2220

tataattttc gtttacgttc tctccctcta ctaatattcc agcagagaac agagcccgcg 2280

ccaggtgtcc agtaccacgc ccctcatatc cgaagctcag gacttggggg ttccgggaga 2340

gagcggctcc agcgcgtcgg gttgtagcta ctgcatctgt gctcttcctt ccccaggaaa 2400

caaattggtgg atcggacctc ccaggtcttt cgcgccccgc caccctctcc cgtgttagca 2460

gggcgcaggg ctccggggcc cctccctgca gtactgggtg atagacccca ctccaccctc 2520

cggttccctc cccccacc acgtgcaggc cagagaaggc aaagaggccc agccaccctc 2580

accagggaaat ttcttttctt tttttgctgg tttcaggctt tttctgcct gagtgaataat 2640

gaaacaaaca cccctgcgc ctcccgcca ccagacacac acgcgcaccg gcactcgcgc 2700

actcgcgcc tcggcctcct agcggccgtg tctggggcgg gaccgctct gcacaaacag 2760

ccgcggggcg ggtggagcgg ggagctcgc gccgcgcgc cagtgcgcgc cggcttctc 2820

gcgcccctgc ccgccacccc ggaggagcac acagcggccg gcgggcccga gcgcaggcgg 2880

cacacccgc ccggcacgc cctgccgagc tcaggagcac gccgcgcgc actgttccct 2940

cagccgagga cgcgcgcggg gggccgggag ccgaggtgtg ggccatccc gagcgcaccc 3000

agcttctgcc gatcaggtgg gtcccgtgg gcgtgcgcg agcccctgga ggccgcgagt 3060

cccgcgcgc ccggggctgc gggcgccgtg gaggcagcgc ggggagagga caggccaccg 3120

cgcggccct gccctgttc tgccctgcc gtgcccgct tttgttctc tcgttacctc 3180

tgtgtcaac tctgacccg tctctgtccc catcttgtc ggcctgagg gctgcgggt 3240

tccacggggt ccgcgggat gaggcgggag aggggaggct cggggcgcgc agaggaggag 3300

gactgcccgc gaagtctcga aaggaggag gggctgtct cccaatgtgg ggcaggggag 3360

gcggaggcct ccctcgcccg ggactaggtg ggaagaggat gcctccgcaa gagggaaacct 3420
gagagtgaag tggggggcac agaaaccctg aacgcacaga gagggagaag tcggggaact 3480
cagagagcgg aggaccgaac ccgaaacccg gccgggggaa actttggaac gccgaaactt 3540
tggcggcgaa aaagccgct gtatcgggtg acaggaagca aagggtcctt cagactttaa 3600
gccacacgtt ccaggaggga gggaggcgcg gagaccgtct gcgggcgcg ctcctcccc 3660
caggaaagac aagagacccg gacggttctt tttgtggtt tgcttgctgt cgtttgcct 3720
cctcttgccc cctgagcggg ccttgctgcc ttgttcttgt gcttggaat ggggtgggtct 3780
cgagcgcgtg gacgtgcggg gaccgggggg gtgggggcga ggaggagtcg gggccgggac 3840
gcctcctagc tggcaaacc tttccaggg agaatccgtt tccacaaacc tgaaatagag 3900
agactgctgg aagtaaggaa atgccaagt cgaagagggt gtgtgtgtgt gtggtggggg 3960
gggatgtgga tgcttt 3976

<210> 8

<211> 8989

<212> DNA

<213> Sus scrofa

<220>

<221> Intron

<222> (1)..(4731)

<220>

<221> misc_feature

<222> (4732)..(4814)

<223> untranslated exon 1A found in some transcripts

<220>

<221> Intron

<222> (4815)..(5241)

<220>

<221> misc_feature

<222> (5242)..(5353)

<223> untranslated exon 2 found in some transcripts

<220>

<221> Intron

<222> (5354)..(8989)

<400> 8

aaaatctgat ttgatctga ttggctagt ttatcacagt ccaccttac ctggtcaaat	60
tcacatactt ctgctgctg cctggctcct gtaggcttcc actcagcatt aattcagcaa	120
atatttactg aacatctgat agatgtcaaa tactgttcca ggtaccagga aagcccagaa	180
gtgaccaaga cagaagacaa gtgctccctc ccacccccca aagagcttgg gttctagtgg	240
aatctgggtc atgacctctt tctgttctg cctccgtag catccccagc ttggtctgac	300
ttcaccacca ccaggggtgt acaaggctga ggtgggacag actcacagaa agacctcaaa	360
cttgtcttcc attccagggc tgctgactca taccatacga ctctgtaagt ttcttccctg	420
atcttcagtt ccctttctta taacttgggg cttgtaatat ttcacctact tagcctctat	480
gttatgtggc ttttgtggat ggcagtgggc tctaaaacggg gcgtgggtgt gaccttgacg	540
gaagatgagc ttatcacgtg ttcaaaaagc agtcttgctt tgaggcaggg agctgactta	600
cctgactttg aggttctctc tgctgaggaa agagtgaagaa cttctgtggg gggtcggggg	660

caaggggtacc ccctggcacc tactgcccac ttgtgaataa ggagcaggtg cctctttctc 720
acctccatct ggggtacttg gcctgaggaa ggggtgagaa ggaccaagag agggtaggaa 780
tagagcgggtt tccttgggtg gggaaatcct ccagtcacct gtgctggtgc tcaagcccag 840
gctgtcatca gtaccgggc ctgccccttc cgtgggagcg cctcacatct cccagctgt 900
caacaaagcc agcttctttc ttctctagga agagtctgac ctatagagct tgaaggactg 960
acatgagccc cagagaggga cttcctggtg tgcaggagga gggctgaggc tcaggatgga 1020
tgcttgacaga ggcaggagtg cttcagcatg gctttggtgg agtctgtcct ggagttacct 1080
ggggcagagg cagatctcaa gatgattagc aatgtactgg cctggaaaga gtcacatga 1140
tttcattttt ccagctcttc tcaaggaaat agacttatag atgcaacctc tcttgactgc 1200
cgtattttat tatgtgggct ttgccaaga tcgtttcagc tctgatactc acaggcgtgt 1260
gtgggggggca gtacttaaca gtaacggaaa cgtcgtgcca ggaacccttc cctccgtacc 1320
tttccccacc tgagggtta catggtcaaa atgactatct gatacacaaa tgtaaactcc 1380
aaggagctgc agcctcggat taatagaaca gcagagacgg acaatgattg agcacctcaa 1440
gcacttttcc gggcgtgtct ccttacttct tgcaatattg ggtaatacgt atctctagac 1500
acttaccatg tgccagctac catccagctg ctgttggtcc cattgtgcag ccgtagaaac 1560
agagacacag agagggttaag cacattgccc aggatcgcat atgggcaggc ctgggactcg 1620
aactccggca gcctgggccc agagtccaca ttcataacca cgggtgtcta ggcccctcac 1680
ccaccccgag cgggtgggat tataattatc ctcaccacac ggaagaggaa accaactaaa 1740
ctgctccatc actcacaagt gacagcaaga atgtcttata cctgccttaa acgtatttag 1800

gattaaaaagt gacagctgca acctttgtat ctgtagcact ttttgccaag aacacttaat 1860
cctccctctc ccacagggcg ggaatccgga cctttgtgtt tctcagctgg aaggggtctg 1920
gggcatgaag ccgggaccct tcacacctgg gctgcagctg ctgagccgca gctccaaggc 1980
cctgcactcc tctgcagggg acatggcaga tggacaggct ctgaatgctg gctgtcatct 2040
gacaggccta tggactgtta gggctggaag gggccttggg gaacattgag tgatgagatt 2100
agtcggcctg gctgggctgg gaaacgtgcc aaactcctac ctggatggcc actggcctcc 2160
tttgatcagc agacctgagg ctcaacttgc acagttccct gcctctccat gaaggaatgg 2220
ccggaagtae atgcttcctt gttttgagag tctgggcata agggatatgc ggagaaggag 2280
gaaggatcatg tcggatcctc tggaagtga attttctgcc ttccaagttt gcatactctg 2340
tcgtgctctg attcatgaac ctggagcctc taattccacg aacctgtagg gtgttcccca 2400
gaggcagctc aggaggaagg gcagcatcag acccaccagc cggcaacttt gagcaagtca 2460
cagaggctcc cagtgcctcc ctccctccc tgacccgggg cgggtgagcc tgaggatttg 2520
ctgagttaaa ggagagaggc tgctttgtaa actggaaggt ggcaaccatg atgggtgctt 2580
gctttttttt gttgttggtg tttgtttttt ttgtcttttt gccttttcta gggccgctcc 2640
tgcagcatat ggaggttccc agcaggctag ggtcaagtt ggagctgtag ctgccagcct 2700
acgccagagc cacagcaacg tgggatctga gccgcgtctg caacctacac cgcagttcac 2760
ggcaaacactg gatccttaac cactgagcg aggccaggga ttggaccgc aacctcatgg 2820
ttcctagtca gatttgtaa cactgagcc tcgatgggaa ctctgggtg cttgcttctt 2880
gaaaggacca gtttatctta gccagttcc tgagcctcca aatgctgtga actttccctc 2940

ccagttgacc acagtcacgc tgccctgcac atttaatgtg aaagatcttc cctgagtcgc 3000

tacttaggtg ctctgtggtg cttggtattg gggcggtgaa cccaagagaa ggaaaaaacg 3060

gggtctatcc acgaccctgt ggcctgaga ccctgtagac tcaggggaag tcagaattcc 3120

caagagaagg cagcttcacg caggaagatt tctgtgcac tttgttttta acacacacac 3180

tgaaaggga tgttgtgag gcattttccc aaggtggaca cacctgcata accactacct 3240

ggctcgagaa acaacatgac aagccccccc cctccccca gcagctctct ggcctcccc 3300

ttcccagtct ctaccactcc cactctgact tctggcacca cagattgggt ttgtcttttt 3360

tttttttttg tcttttttag gctacacttg gggcatatgg aagttcccag gctaggggtc 3420

caattggagc tgtggctgtt ggcctacacc acagccacag caacatggga tccgagccgc 3480

atctgcaacc tacaccacag ctggtggcaa tactggatcc ttaaccact gagtgaggcc 3540

agggatcgaa cttgcattct cgtacatact ggtcagattt gtttctgctg agccaccatg 3600

ggaactccct ggttttgtct attttttttt tttttttgt cttttttgcc atttcttggg 3660

ccgctcttgc ggcataatga ggttcccagg ctaagggtcc aatcggagcc gtagccccag 3720

cctacgccag agccacagca acgtgggac cgagccgagt ctgcaacct caccacagct 3780

cgcggaacg ccagatccct taaccactg agcaaggcca gggaccgaac ccgcaacctc 3840

atggttctta gtggattcg ttaaccactg cgccacgacg ggaactccc gttttgtcta 3900

ttttgaacg ttaaataaat gcaagcatcc agggctgctt tgactcagta ccatgtgtga 3960

gatttaccct gttgatgtca gcagctgtg ctggttcctt ctcacggatg tgtgtgacct 4020

tcacctggac cacacctgat ctggctgatg atgggccttg gggtttttcc agcttttgg 4080

cccaggtcac gtctctgttt gaacttaaat gcacttgctt tcaggtatta atctggggcg 4140
gaatgactgg aacatgaggt gtggttggtt cagctttagt acatgccagc agggaggatt 4200
tcagtagttt attaagcaga tcttgaagac tgtggtcaac tagctcatgc cccacaggag 4260
ggggcggtga atttcttccc cagaacagga gtgacaagct aaattaggca tccatccgct 4320
ggaagctgag ggggcagttc ttggctcctt tctgtcaggt ttcggcccct tctccttagt 4380
ctggggtttc taggctctac tcccaggaag tgtctggggc cacttgggaa caatgggtgg 4440
gggggctctg agcccctact tacttcattt cctccttca gccaaagccc cctgtgtcct 4500
ctgttttaca tagtgggggt ctgagaatga cttcattttt tttttttttt tttttaaagc 4560
tttagctgtt gcgacattta caaatccact gctgtgaggt ctcttccagg taggaaattg 4620
tattttggga gcaggaggtg ggtgtgggga gggtaagca ttattcagcc aaagagttgg 4680
gttgggcctc agtgacctt tgaagttctt atagcttggc ttgccatgca ggagatctca 4740
gaacattcta taaaaatagt gttcaaacag aacaacttct gaagcctaaa ggatgcgaac 4800
aagaggctcg gaaggtagca tttcaacggg agttttgagg atgctctcct ttagccaccc 4860
ctctccattt tctgcccctt tctttttaaa ttctccattg gctgtccctg ctagtgtgca 4920
tttgggggtg tttgggttca gaatggttct cattttcgcc gaggagtggg tgatgtgggc 4980
ggcctgtgtg tctctcccaa ggggtgtggc tgtccctcct ccaccaccag gcctagtttg 5040
gacctgtagt ttcgcttagt gaaggaggcc gggccgatcc tgggccggag agagacgtct 5100
ctgccttggc atgcagctct gagtcaacag gcctgataaa cagcccactt cccagggcga 5160
gcaaggagga acaaggcccc tggctgctgt gggatccgtc tgcgtcctc ttcgtgaaac 5220

cgctgtttat tcttttgaca ggagttggaa cgcagcacct tcccttcctc ccagccctgc 5280

ctccttctgc agagcagagc tcactagaac ttgtttcgcc ttttactctg gggggagaga 5340

agcagaggat gaggtacgtg aaacgttgaa atgatttacc tccgctttgc tggggtcacc 5400

gggggggtgg gtatcatgag ctggctgcag cgtggagaga ggagcccccc tctccccctg 5460

acttcttgct gctcccccca gttgttctga aagaagacaa agtcctccag tccccggcat 5520

cggatctagg agtgggagct ggcaggatgc tggctcagtc actgttggtt ctgctttcgt 5580

tggctgcccc gcaggacctc acgggggtgtg gctacagcct ggggttctct gtgtgggcca 5640

cacagtgccca ttgtggggcc aggaggacga gtctcaggcc cgggacctgt gctggggggc 5700

gacatagtgc cctctcaggg cagcaccgat ccttcatgta cctcgcccta tttctcttgg 5760

aaaaactctt gcaccatgat ttctgagcca ggcagcaagg agaagctggc tggatccagg 5820

cttcagattt ttgaagggga ttcaagaaag gggcctacaa gatgtccctc cgagaacagg 5880

tctgtgatgg ctggagcgac agctgtgaaa aaaataagtg gaaagagcct tcggtgcggt 5940

actccccccc caccctgcc ccccaaatta taccatgttt cttccaacag ggagcatttc 6000

cctgtaatgc aagccaattt aaattcttga ggggtcacat tttggtttta tttcaactga 6060

ttattagtgt agaggagtat aagataacat ttctttaaaa accatcaaca caaacccatc 6120

actcgtgatt caattgttta ggagaggagg gaactccgcc tcgtatacca aatacagtct 6180

gctctcggtg cagcgtgcag tcccagcaag gccctctcct cgaactcaca cagctcttgt 6240

ctccagcggc ttccttccca tgtcttggct aggtctgggt ttcttagtaa ccccaaaggc 6300

ggagaatcaa attcacagat ttttttttcc tggatattta gatcttgtat tttaagccac 6360

actatttata aggctcagag atacatttaa actctgacta gggcttctta taaaagtgat 6420
atctggaaag aaggtctggc tttaacagag taagggtcag acccccctt ttcccattaa 6480
tgactccagg aatgctctgg aagactgaag tggaggcaaa gaaggacttg aatttgcattg 6540
acctgatctt gaatccaggc taaatttttc ctggctgtgc gccttttaggt gggtcattta 6600
cctcccctaa ttctcagggtg gctcacttca tcctctattc ttttactgag gcagagaggt 6660
ccctctacca ccagggttgaa tgagctcagt gacctctgaa aactccaaag tgctgcacag 6720
atcaagggtg tatgaggtag aagaggaagg gaaaaaggaa tgagtaggat caaagaaaga 6780
aggagtgaag agaagcagag tggagagaca gagccaacac aaggatctgg gtaccacttc 6840
tggattaggg tcagggttga gaagatgaca ttgatggttg ggtctttttc actacacaga 6900
gaatagagct gaccattaga cttggcccgg agccagtcct tgtgaaagaa atcaatattc 6960
agattatcat gacaactacc atttgtgtaa ttttaattca caggatcact ttttctggcc 7020
cacgaggttg aaataagaat ggctggtcag attgactggg gcggtccgac tggcctgtgc 7080
ttgagagttg accatgagct ccctgccatc tagcgtgtat gtcaccaga cttttaactc 7140
accatctgga ctgaccctcg agaacttgat gccatttgag agcacccaag gggtcagag 7200
gaccttatca aatcctctga ctccctctgtg caggctgttg gccagcttat actccttccc 7260
atccaacgtg atgttccttt ggcaatttgc tttgccacc tgccaaccac tgctccaaag 7320
tagggatgct tttggaggta cccttccaat tcagcaaagc caagcaccac atctgaggct 7380
ctgccttgcc tgtctttgac ctccagggcc gtgatggtgc agcccagga gatgatttcc 7440
actcccagtg ttgttcagcc cgaggagatg atttccaatt ccagttggt ctgcttgag 7500

ctggaatttt tccatgttcc ttgccccaa ggggagttct ccaaacacag atcttgtaac 7560
tgaaaccatg aggaaagctt ggggtgtgta ggtgctccag gtccttcaaa cgccccatct 7620
tttggcagtt tcttgctcag gtgggtccag ccagagtcct ggagaattca gctctttgat 7680
cctggctgga gtggggggtg caccaccagg tgattgtgag gtctggatcg tgacctgtga 7740
gcaggagacc aagtagcatc atgttcagct ctttctcctt gggatcaaag tgagaggctc 7800
caaggagctc agcaaggtct acctggatgg ggcaggttgc tcctaggacc caggtaggtg 7860
cggggagcag ggtcagtacc tgggctccac ctgcagcccc aggacaggca cccaggctgg 7920
aacgattccc ccaggcaggg gcagcacctc acctggagga agcatttggg ccttgcccac 7980
tccacacccc aggctgcct gggggcctga cccggaggct tctgggtgaa gtggcctgag 8040
ggctcaacac attttgtggg caatcctatc tcttttttta tttttatatt tttattttt 8100
gctttttagg gccgtaccg ctgcatatag aagtttctctg gctaggggtc aaatcggagc 8160
tacagctgcc agcctacacc acagccacag caacacagga tccaagccgc gtctgtgacc 8220
tacaccacag ctcatggcaa tgccgatcc ttaaccact gagcgaggcc agggatcgaa 8280
cccgaacct catggttct agtcagattc atttccgttg cgtcatgacg gaaactctgg 8340
caatcctatc ttttgatcac cacttctagg aatctgtggc cactgcagca agttgagctc 8400
cagtgaacct gtctcataa aaggagcctt cagctctgtg gctgccttct catacaggctc 8460
ttggctcatt caggggaagt taagcccaca ggacatgttt caaaggacgg gaaatgcact 8520
gggttttagc acagtctgca cgaggcccgg gagtgggggt gcaagtgggt tcttttgaa 8580
accgctgcag gggctgagtt gtgggagtg cccaggagca gagagaaatg gcaaacgcct 8640

21

tggcaggagg gcctgtggga tgggtgggagg gctcagggtg aactgggccc gctgggttca 8700
cctgatcctc tgagggctgg ggcccagggtg gtgctgaggt ggttacctc tcccttataa 8760
gacaggatgc tagtgcctc taggtcttaa tcctgtgctc tccctcttcc atgagaaatg 8820
tagaagcaac ccccaactttt cctatttggt gggttaagata gtcaaccacc aatcttgaga 8880
attagagagt ttgaaaatt ctgtgacaaa cacatccgtg aagggtttt agaccacatg 8940
ggctgccaaa tgcctcattt taatccagag agaaaaataa aattgtttt 8989

<210> 9

<211> 240

<212> DNA

<213> Sus scrofa

<220>

<221> Intron

<222> (1)..(29)

<220>

<221> misc_feature

<222> (30)..(118)

<220>

<221> Intron

<222> (119)..(240)

<220>

<221> 5'UTR

<222> (30)..(38)

<220>

<221> misc_feature

<222> (39)..(41)

<223> This "atg" is the translation start codon

<400> 9

aatttttcct tctccttttc ttttcccagg agaaaataat gaatgtcaaa ggaagagtgg 60

ttctgtcaat gctgcttgtc tcaactgtaa tggttggtt ttgggaatac atcaacaggt 120

aattatgaaa catgatgaaa tgatgttgat gaaagtctcc tctaattctcc tagttatcag 180

ccaagtcacc agcttgcatc aaaagtagga ttcactgaca ccgtaaagaa agcattccag 240

<210> 10

<211> 2685

<212> DNA

<213> Sus scrofa

<220>

<221> Intron

<222> (1)..(2140)

<220>

<221> misc_feature

<222> (2141)..(2176)

<223> This region defines exon 5

<220>

<221> Intron

<222> (2177)..(2685)

<400> 10

aagcttttaa ggactctaag ccttcatttt tctttttttt tttcctatct tcgacttggt 60

tgctaggaag cttagagcaa agtattgtgc ttaaagtgtt gcattttcct tggccttcac 120

tttttttaaa acattttttc ttattaaagt atagctgatt tatagtagcc ttcattctgat 180

atgatttata ccctggtggt aaatcctggc ttttggtaga tgccatggga tcttggaat 240
ttgctcaaac tcattttgcc aatatcttag ctatgaagta aaaataaagt taaagatttt 300
gttctcacag agtggctggg atgaccaaag tcatgtgaaa acacccgagt gactaaaatg 360
tttctctggt tcggtttggt ttgttttgat tcttgatttg ttttctatt tctgtaacc 420
acactttctt cataagccat ttcaagcact tctgaaagt agatggactt taagtttctt 480
ggacttccag ttgtggcgca gtgcaaaca atctgactag tatccatgag gatgcatctt 540
cgatccctgg ccttgctcag tgggttaagg atctggtgct gctgtgacct gtggtgtagg 600
tcacagaggc ggctcagatt ccaagttgct gtggtgtgg cgtaggccgg cagctacagc 660
tccaattaga cccctagcct gggaacttcc acatgccgca ggtgcaacc ccaaaagata 720
aatgaataaa taaataaata tgcgaccttc ctttcttggg gcccttgcac gtttttctct 780
ctgttaggca cactcttgct aatccctctt cactgggcct cctatgtatc cttcagaact 840
cagctaaaac atcatcccct cccctgggga gccttcgagg tcttctggtt aagtgtcct 900
atgctttctt ggagttttga agtcctataa tgatgtgttt atcaaaatag ggtccaccct 960
ccctgccagc ttctttacac cacagacaca tgggtgtctgt ttcagtcaac actgtatgtc 1020
tggcacttga catgtaacgc atgctcagca ggtatttggt gaatgaatgg aggcggtctg 1080
ctagagtcgt catatattta ctgatccgt cttgtaggat ggtctcactg cttttgtag 1140
cttaagaagt acctttttt ttttttttt tttaatggcc acacccatgg catatagaaa 1200
ttccacgaag gaaggaagaa agaaagaaag aaagaaggaa attcctgggt cagggattga 1260
atccaagcca caggtgcaac ctgagctgca gttgcggcaa caccacatct ttttaaccac 1320

tgtgctgggc cagggatcat acctgtgcat ctacagcgac ccaagccacg gcagtcagat 1380
tctttttctg cctttctttc tttcttttct tttttttttt tttttttttt tttgtctttt 1440
tgccttttct aggtgcgga tatggagggt cccaggctag gtgtcgaatc agagctgtag 1500
acgccggcct aaaccacggc cacagcaaca caggatccaa gccttgtctg tgacctacac 1560
cacagctcaa cggcaacggt ggatccttaa cccgttgagc gaggccaggg attgaacccg 1620
caacctcatg gttcttagtt ggattcgta accactgagc catgatggga actcctgcag 1680
tcagattctt aaccacccat gccacagcag gaactcctag aagtgccctt tgaggctact 1740
ctgtagacag ctttgagcca gcgaggcaag acctgttttt ctggaggaag ataaatcctg 1800
ggtagaggat gggtaggctg tggcttctct gggaccatc tctggagcct ctctccctca 1860
gcaaagccac cttggacaat aagagctgcc atctattttt tttttcttta aactaagatt 1920
tgatattttc cagagacctc cctcccaccg ttgatctga gtaattctga aatgacgaga 1980
gccccgtgat atcatttttt cgatctcgaa ggtggaaacc tgggagtagc cacaaccag 2040
gctctcagct cagcctaggg tttcaatgat aatgattgca aaatagcttt tctctgcgtt 2100
ccaagtaaca tgatatgttt ttatttccat ttgcttttag cccagaagggt tctttgttct 2160
ggatatacca gtcaaagtaa gtgctttgaa ttccaaatat ctctagggtca ccttccatgt 2220
gaccctggtg gccctacagt ccattcttaa catggcaggt ggtgacgcac ttgtggtcct 2280
aggtaggagga gagggatggg gttccagggg tctgagctgt acttctccag ccctagact 2340
tgcctttcta gagcatgagt tgtgtttttc ctttgcttct catcaagtat ctatctcttt 2400
aagtgatgtt gtttgagaaa cattcctgcc ttgctcataa aaaagaatca gagtagatat 2460

25

tatccattat gctacctact acatgtgga taaagaccct tgcccagaaa ttttgccaag 2520
acaaaggatt aggaagaaag gctgggtgtc ctgataaact aagtgtgtgt attattatta 2580
tttaatatta ttactaatac tgggtgattt aagggactcc taaggccttc aatttttcoct 2640
tttttctttt tttttcccta atcttccgac ctttggtttg cctaa 2685

<210> 11

<211> 180

<212> DNA

<213> Sus scrofa

<220>

<221> Intron

<222> (1)..(37)

<220>

<221> misc_feature

<222> (38)..(100)

<223> This region defines exon 6

<220>

<221> Intron

<222> (101)..(180)

<400> 11

tttctaaaaa atgtttgtca tctttttcat ttcttagaaa cccagaagtt ggcagcagtg 60

ctcagagggg ctggtggttt ccgagctggt ttaacaatgg gtaagactgg gaaacggcca 120

tctgtgtatc tgctcaaggc tgtagagtcc aaataaaatg gtttcacagc catgaccttc 180

<210> 12

<211> 242

26

<212> DNA

<213> Sus scrofa

<220>

<221> Intron

<222> (1)..(100)

<220>

<221> misc_feature

<222> (101)..(205)

<223> This region defines exon 7

<220>

<221> Intron

<222> (206)..(242)

<400> 12

atgaccttct ccagtcgcgt cgtccttctg gcttattgga cattctggca catgggtcac 60

cctccctgcc ttccctcagct tgttttccgt ttgtacgtag gactcacagt taccacgaag 120

aagaagacgc tataggcaac gaaaaggaac aaagaaaaga agacaacaga ggagagcttc 180

cgctagtgga ctggtttaat cctgagtaag aaaagaagcg ttgccctatt tcagtaaata 240

ca 242

<210> 13

<211> 720

<212> DNA

<213> Sus scrofa

<220>

<221> Intron

<222> (1)..(257)

<220>

<221> misc_feature

<222> (258)..(394)

<223> This region defines exon 8

<220>

<221> Intron

<222> (395)..(720)

<400> 13

agcagaacag ggggacggaa gtacatacac gttgtacagg tacgatcccc aaagggccac 60

cagggcagcc cgcagaggca cttgggccag agcctcctgt ccttccccca gaagatgccg 120

caatgtcaca ccaccagctg actggggcta aaatacagtc aggattcaag gccagtcccc 180

caagccatga ctgacccatg ttcccccaaga ctgtcgtacc ttagcaaagc catcctgact 240

ctatgttttg tcaccaggaa acgcccagag gtcgtgacca taaccagatg gaaggctcca 300

gtggtatggg aaggcactta caacagacgt cttagataat tattatgcca aacagaaaat 360

taccgtgggc ttgacggttt ttgctgctcg aaggtagggtg ttgctaataa aactggcctt 420

gagtttttcc ccttccacta tcagaggatg ggtgaggggc ccctggggtt acagaggctg 480

ttcatgtcat gtctgaatta gtggagagga gaatggtgtc acagggccat ttagactcc 540

cttctgctga ggtccccaaa ggctaagaat aaaactagtc agagggtcaa ctctttccca 600

cctcaggggtg aggggcttgg gttgcaggga agaaaatctg ctatacccac tgcacccaaa 660

gtcgacagta cacccacagc cacctccacc ctgacctcca cgccctctg tggaaattcc 720

<210> 14

<211> 2964

<212> DNA

<213> Sus scrofa

<220>

<221> Intron

<222> (1)..(318)

<220>

<221> misc_feature

<222> (319)..(2904)

<223> This region defines exon 9

<220>

<221> terminator

<222> (1010)..(1012)

<223> This is the translation termination signal

<220>

<221> 3'UTR

<222> (1012)..(2964)

<220>

<221> terminator

<222> (2858)..(2863)

<223> This is one transcription termination signal

<220>

<221> terminator

<222> (2883)..(2888)

<223> This is one transcription termination signal

<220>

<221> polyA_signal

<222> (2904)..(2904)

<400> 14

tgcaatgccc agagcagctg aaaacacatg ttctctctgc ctggttggt tccaagagtg	60
agagaggaag gagcagggt gagcatgccc agccaccctg ccagaatcac cagtçaggta	120
agccactcca cctcccaaaa gctgaatgac tgaatggtg agagtagctg ggaatgttac	180
agcaacagac gtctctcatc caggatggg aaaaatcatt ctttcctaa actgcaaaat	240
acagactaga tgataatagc atattgtctc ctctagaaat ccçagaggtt acatttacc	300
cattcttctt tatttcagat acattgagca ttacttgag gagttcttaa tatctgcaa	360
tacatacttc atggttggtg acaaagtcatt cttttacatc atggtggatg atatctccag	420
gatgcctttg atagagctgg gtcctctgçg ttçctttaaa gtgtttgaga tcaagtccga	480
gaagaggtgg caagacatca gcatgatgçg catgaagacc atcggggagc acatctgtgç	540
ccacatccag cagcaggtgg acttcctctt ctgcatggac gtggatcagg tcttccaaaa	600
caactttggg gtggagacc tgggçcagtc ggtggctcag ctacaggçct ggtggtacaa	660
ggcacatçct gacgagttca cctacgagag gçggaaggag tccgagçct acattççgtt	720
tggçcagggg gatttttatt accacgçagc catttttggg ggaacaccca ctçaggttct	780
aaacatçact caggagtçt tcaagggaat cctccaggac aaggaaaaatg acatagaagc	840
çgagtggcat gatgaaagcc atçtaaacia gtatttcçtt ctcaaciaac cçactaaaat	900
çttatcccca gaatactgct gggattatca tataggçatg tctgtggata ttaggattgt	960
caagatagct tggçagaaaa aagagtataa tttggttaga aataacatct gactttaaat	1020
tgtgççagca gttttctgaa tttgaaagag tattactctg gctacttctc çagagaagta	1080

gcacctaatt ttaactttta aaaaaatact aacaaaatac caacacagta agtacatatt 1140
attcttcctt gcaactttga gccttgtcaa atgggggaat gactctgtgg taatcagatg 1200
taaattccca atgatttctt atctgttctg gggtgagggg gtatatacta ttaactgaac 1260
caaaaaaaaa attgtcatag gcaaagaaaa agtcagagac actctacatg tcatactgga 1320
gaaaagtatg caaaggggaag tgtttgcaa caaataaga ttgggagggg tcgtcctctt 1380
gatttttagcg tcttctgtc tctgctaagt ctaaagcaac agagttgctt tgcagcagga 1440
gatcagagtc taccttagca atcctcagat gatttcaaca gcagaggact tcaggttatt 1500
tgaagtccat gtccttttcg catcagggtt ttgtttggct tctgcgcagg atactgatca 1560
agattcccaa tgtgaatgtt ggagttacag ggaatccgaa tgaaccaatg ggagctcagc 1620
acgaaataaa agcacagctt ctaagtaagt ttgccatgaa gtagcgaaga cagattggaa 1680
agagaggggg ctgatcactg tggggcaatg ccatttctaa gagacacagg gcatggagtt 1740
ggcatgtaca tacagcttgg atccaggcac tgaatgggag gcaatgagag tggctccagc 1800
ctcctcaacc atatgacaac tagagcagca ctgtcttaga agatgcttct tgctttggcc 1860
aagtcattatt cagtctgcca gactctggaa cttgtgtcta caaatccttg ctgagaggaa 1920
gtggatgatg tcagagtgga cagaggccta cattggggtg aagtgacttc ctagaccttg 1980
gcttcatgac aatcaggcat cagcaagccc tgctgccacc tgctctaact ctcagagtcc 2040
ctcagcccat catgggcaac ttgagagcca ccgtcaagga gtggactaga ggaaaagcct 2100
gcttatcagg gaacctctca ttccccctgc ccagctgca ctactgaagt gtaactgccg 2160
gacatgttta ataaagtggg taattgattt tatatcaaag tagagaggat ggcaatggga 2220

gacccagttcc tcatgactaa acagcttttc aatccctttc tctaagaaaa gctatgagat 2280
cttacatgta atttaaagtt aagcagtttg gtgtaaagga agttaggagg caatatttac 2340
atctgcaggt atgtgatata cttttgcttg tgttccagtt taggtcattt gtgtccattt 2400
tcaaatgatt tacttgaaga gccattgcac tgacttgatg ttcagcacga tgggcttctt 2460
tgataaaatg aaacctacat tttctctact gtttccctgg gctcctact cttcaattct 2520
tgctaaaaat ttttgcaacc cagcaaaata actcaacaaa ataaccaaac aaaataactc 2580
aacaaaaatc ctggagaagt agtcttgtaa aagaaaaagg aaatcacaag tcaattagga 2640
ctcttgtttc tctataacgc aagtttatgg aatccattct ggagtgcaga gacttcatgg 2700
tgcaagttcc aaactacaga aatgattcgt tctcaaagat taaagaaaag gactgatatt 2760
tccttttgaa ggaatcttga tttttaaaaa aaaaatcatt taaatttaaa tttcaaatgg 2820
acaaattcaa gatcttatta atagttcaat attaaaaaat aaaaattcct gatttaaaat 2880
taaataaatt attttctcag tatattctgg tctggtcatg gattgtggct tttttcccaa 2940
agatgttcag aactgtcatt taca 2964

<210> 15

<211> 1500

<212> DNA

<213> Sus scrofa

<220>

<221> misc_feature

<222> (1)..(1500)

<223> genomic sequence between exons 2 and 4

<400> 15

ggatccttaa gccactgagc aaggccaggg atggaacca caacctcatg tttcctagtc	60
agattcgta accacagagc cagcagggga actccacac attatttatt gacggccttc	120
tctgctctct gtggggcact gggaattcag gggatgacaa gaagtcaccc ctctgccct	180
caggaagctc aaaccactca ttatttattg acggccttct catgctctct gtggggcact	240
gggaattcag gggatgacaa gaagtcaccc ctctgccct caggaagctc aaacaagcag	300
gtagaggagg cagagcaaaa tgcaggctt atccggtag ccgactcca gggcgatgtg	360
tacagcaaag gaatagagg atgggggccc gaggagagaa aagggttca gccgtggtca	420
gggtgggggt gggaagtggc ttcacaaagg cagtgcatt ggctcccagg tgtccactct	480
tctgtctctg ctacctctg gtctctcct tgtgggccct cctctatcct acctctaaag	540
cttcagcca gcacctcct cctcctctt ctctctctgc attctctcct gggtaatcaa	600
attcgttccc ttcacgtcag atccggtatc ttcttggtc catgaacaac ttctccgatt	660
gcacggtctg cctacatctc tctgatgaac ttagacttg aatgtccact tgtctccctg	720
ttccctttta ggtattcgca cactccccga cattcacag tccaaaagg aattcatgat	780
tattatctc caagcctgtt cctcctccag cccatctgag aaaatactac aacccccctg	840
cttaagcaga aatcttgggt ctccctgtc tcctctctga taacaaaatt accaaccag	900
tcctatcaat tctctctcca aagtatatat atatatttt ttttaattt tttccgctg	960
tacagcatgg gcatcaagtt attcttacat gtatatttt ccccccact ttgttccgtt	1020
gcaatatgag tatctagaca tagttctcaa tgctactcag caggatctcc ttgtaaatct	1080
aagtgtatc tgataacccc aagctccga tccctccac tccctccctc tctgtcggg	1140

cagccacaag tctattctcc aagtccatga ttttcttttc tgtggggatg gtcatttgtg 1200
ctggatatta gattccagtt ataagtgata tcatatggta ttgtcaaag tatatatattt 1260
atttttcttt gtctttttgt cttttgtctt ttttttggtg ttgttggtgt tgcgttggt 1320
gttggtgcta ttacttgggc cgctcccgcg gcatatggag gttcccaggc taggagttga 1380
atcggagctg tagccaccgg cctacgccag agccacagca acgcgggatt cgagccgcgt 1440
cggcaacctc cacacagctc acggcaacgc tggattctta acccactgag caagggcagg 1500

<210> 16

<211> 500

<212> DNA

<213> Sus scrofa

<220>

<221> misc_feature

<222> (1)..(500)

<223> genomic sequence about 4-5 kbp downstream from porcine exon 4.

<400> 16

ggtacccatg aaaagcccaa caacacaggc tagaaggagg atgtcagaga gagagagcaa 60
aggaacgtga gagttcaggg agggcaaggt tatgtttggc ttggagatgg atctatgttt 120
tgcatttatt tttttggggg ggggtctttt tgctacttct tgggctgctc ccgaggcata 180
tggaggttcc caggctaggg gtctaattgg agccgcagcc accagcctat gccagagcca 240
cagcaacgca ggatctgagc cacgtctgca accttcacca cagctcacgg caacgccaga 300
tcgttaaccc actgagcaag ggcagggacc gaacctgcaa cctcatgggt cctagtcaga 360

34

ttcgtaagc actgcgccac gacgggaact cctcattta gaaatattta ttgagcacct 420

actgtatgcc aggcattggt ctaggttcat accaaagaag gctcaaaaag atggcatccg 480

aactgttgcc cttgaaagga 500

<210> 17

<211> 1520

<212> DNA

<213> Mus musculus

<220>

<221> Intron

<222> (1)..(1130)

<220>

<221> promoter

<222> (381)..(1321)

<223> Fragments and derivatives have promoter activity.

<220>

<221> 5'UTR

<222> (1131)..(1320)

<223> untranslated exon 1

<220>

<221> Intron

<222> (1321)..(1520)

<400> 17

tcccaatgca tcttttccca gtgggtcttt ggattcatgc tgccatatga tctgctgata 60

ccatgcttca gtaccaagtt gattctttgc tcttgcctg atgctgaaga cctaaaatga 120

tgaaaatgga aaaagaatga agaataagta tacacacacc cggcctgctt ttgcggatca 180

ggtgggtccc gccgggcgtc tgacactgaa gccacgcggg ggcttcagtg gggaggagggt 240
gtgggagagc gcgagcgccg ctattccggc ccagccctac ctcggtcctt gcttttgtcc 300
tggtcactcg atcatttcct ctgtatccac ttctgaactc taggctctgt cccaccctga 360
acagtgtcgc tgcattctgtt tgcttactgg ggtctccgc caccttcctt cgctatccga 420
atagctgata ttcagggcag cacagggcag ggcagggcag ggcagggcga gtagggcaga 480
tcagatcctg ggaccaccgg tactaaccag tgagtgtaga aagcaggagg tgtcttttcc 540
tactgtagtt aggacagggc gggttggtc ttcttatgga caagatggaa aaggggtgca 600
ggtaggggca aagtgagaga cactcgaatt tgagagacag acagactcct aacagtgaag 660
gaaggaccaa gccaaaatca agcctgggca aagtctcagg cactaacttt gctgtgttgg 720
gtgatgggag gtaatctcgt cacaactttt caaacacact cgttcccact gcaaggagac 780
accatcaagt gtttgaagat ggcaggggaa cctctcaaca aaacacacac acaaacgttt 840
tattatttta tatttatttt gcatgcaaag tactgtgttt cattatggca tttcataca 900
tatgcgattg cacaaactct tgaaaatcat ccaagaaaca gcaaagcggg aaataatgtt 960
gtgggggggg gccgcggagg agagagaaca gagactggag agagtgtgtt cctccttgct 1020
gcggggggcca ggaagaggct aggagggcgg ggatgtcaac gccactagct cctccctcag 1080
gaaggacccc agggactctt atttttgtag ttttgttgt ctgggccact atcgggccca 1140
gaacagatct gactgcctct ttcattcgcc cggaggtaga taggtgtgtc ttaggaggct 1200
ggagattctg ggtggagccc tagccctgcc tttcttagc tggctgacac cttcccttgt 1260
agactcttct tggaatgaga agtaccgatt ctgctgaaga cctcgcgtc tcaggctctg 1320

ggtaggcaaa ggcgaggggg ctcgccatgg ctcgggttgt ccagggattg gggcatcagg 1380
actacgggag tctctgcctt ttgatagtgc ttccttacag ttatttttgg gagtagttgc 1440
ttcttcctga tgggagccgc gtgcgggtcc aagctatctt ttgcaagtaa caggtgtctg 1500
nnnnnnnnnn nnnnnnnnnn 1520

<210> 18

<211> 1207

<212> DNA

<213> Mus musculus

<220>

<221> Intron

<222> (1)..(653)

<220>

<221> 5'UTR

<222> (654)..(773)

<223> untranslated exon 2

<220>

<221> Intron

<222> (774)..(1207)

<400> 18

agccctaggt tgtcgttggc tacacagtga gttcataggc tgctagggat cctatctcaa 60
aaaggaaaac aaacaaacaa acaaaggggtg ggcagggtta agccttgtcc ctcaggagca 120
ggtatgggtt tctgaggctg tcccaagtgc atatggtaaa ggcttctcta tggagattta 180
caccattttc taaagtgcag tgttccacat aactgtgtgg cttccagagc caggctgtgg 240

37

aggaagagct tatctcagaa ccacattttg gcgtcccatc aaagtgcctt gtccgctaac 300

ctgcctctgc cccaggctgt gtcatacgca tctccgggga ggcatacactt gagaatgagt 360

gcatttcaca gggctcccag tttcccttg ggactgggtg atgtggaggg tggaggcctc 420

atcgcttggtg actcctggca tggtttggtc ctgcagtttt tctctgggt gaggaagtca 480

gaggaccaac ccagagccct gattctgcct tgctgcgtag acctgaatca acagccctga 540

taaacagccc atttccggg gctgaggga caaagcctgt ggctgctgcc gagggatctg 600

tctgcccacc ccaccctoc tcttctgaa acagctgttt attattttga caggagtgg 660

aaccctgtac cttcctttcc tctgctgagc cctgcctect taggcaggcc agagctcgac 720

agaagctcgg ttgctttgct gtttgctttg gagggaacac agctgacgat gaggtatgtt 780

taaaggattt gtgtctccca gccttgggtc actgcgagct actgttaggt caccaaattg 840

ttccacctga gggaggaccc ttgtctctt ccgaagcttt ccttgggtccc ttctgtgatt 900

tgtgtctctt tcccttttgt ttctgaaaca ggggctgggtg gaatgctggc tggggacttt 960

ggtattctgc ttctcttggc agccccggg gctatgccag tcaaggctgc agcctggagt 1020

tctctgtgtg gggtttggtt tggcggggct gagtcttggg cagggcgagg tgggagggtg 1080

ctgagtcttc tctgctctgg gctgtctcgt acatgtccgt tggctggctg ttctgggag 1140

gtatcacttg agattgattg cattccacat gacactgctc ccagggacag cccggcactc 1200

nnnnnnn 1207

<210> 19

<211> 900

<212> DNA

<213> Mus musculus

<220>

<221> Intron

<222> (1)..(336)

<220>

<221> misc_feature

<222> (337)..(517)

<223> untranslated exon 3

<220>

<221> Intron

<222> (518)..(900)

<400> 19

ttccggcatc tttaagatct tgatgtaccc aagtcaactt cagcttcaca gcttcttgtt	60
tcaatgtctg ggatccacac ctgatcttct gggatcctcc aaagggcttg ggtcacttct	120
ccatctctgc cctctgtagt actctaggct ctagttgact ccactccact gctgctgctg	180
ttcttggtga tcacccatg gtactggcaa gtaggtgaaa gaagaagagt gaatattcct	240
tcaccaatg tccttatgta ggccctcagc agaagggtg gctcagatta aagggtgcta	300
cccccatgcc tggatctaaa acttgctttg ttccaggctg actttgaact caagagatct	360
gcttacccca gtctcctgga attaaaggcc tgtactacat ttgcctggac ctaagatttt	420
catgatcact atgcttcaag atctccatgt caacaagatc tccatgtcaa gatccaagtc	480
agaaacaagt ctccatcct caagatctgg atcacagggtg tgcccttctg tttctggatt	540
atagttcatc ccagatgtag tcaagttgac cactaggaat agccatcaca agcccgttgt	600
ggaggctgcc cctgcccc cgccccgcgc gccctgagg ctctcacccc tttcttgtgg	660

cagctcttgt cttcatctcc agtgatcaac tgcattccc actctgcac ttgccttct 720
gaacaatcac cgtcccaaag ttcttctcag tcttttgta tcctcttccc tttctttcac 780
aatcttatgc agaatttaaa aaataacctga ctccttcagt agttccagtt gtttgctggc 840
ttgtgggggg tgtagtgagg tgaaccaggt ggggctgaaa agtgggtgca nnnnnnnnnn 900

<210> 20

<211> 1020

<212> DNA

<213> Mus musculus

<220>

<221> Intron

<222> (1)..(479)

<220>

<221> misc_feature

<222> (480)..(568)

<220>

<221> 5'UTR

<222> (480)..(489)

<220>

<221> misc_feature

<222> (490)..(492)

<223> This "atg" is the translation initiation codon

<220>

<221> Intron

<222> (569)..(1020)

<400> 20

40

tctctttcca atgccacat ggatgggctt cagcatcctt tcagatcatg aagcctcatt 60
aactgtgctg gcctaattgg ccatgactag tttgtgtgct tgagggatag ggggagggga 120
gacacttgct gctgagttag ttacaaatgt atcctgttag gaaggatgtg ggcagatgcc 180
tttcattatc ttactgcat caaacatttt atgggtatga gtgttttgcc tgcaagtatg 240
tatatgtacc acttgtatat gtggacccca tggaggccag aagagcatca ggtcctgtga 300
aaccagagtt atggacacct gtgagctgca aatgtggatg ctgggaactg aatcgagcag 360
gtgtttcatt gaggtgtttc aaccacacag ctgtttctcc agccccagaa gccatctctc 420
attccagatt tagtttattt aatctatttc cccctctttt tttctcctg cctctacagg 480
agaaaataat gaatgtcaag ggaaaagtaa tcctgttgat gctgattgtc tcaaccgtgg 540
ttgtcgtgtt ttgggaatat gtcaacaggt aattatgaag ccagctagaa aggctgcttt 600
cattccctgt gactggtgcc agctgagtga ccaatcagtc tgaacataag ggacggagcc 660
gtgagcagga gtccagtctt cctgtgttcc tgagccccag atggccatta aaactgtaga 720
ccatccaagt cacttctgcc ttagtaatta tcctctttca tgccgtgctc ctcaaacctc 780
gaatttctgt aagctagatg gagagagaaa gtacattaag ccaaaccac catctcaagt 840
aatttgata agcagatccc agaagattca ggccaggcag ggtagtgcac gtatggagtc 900
cttgtgcttg caaggcagag gcaggagcat catacaaatg gaagaccaag cttgtcttca 960
tagtgacttc caggccagct gtagccttac aaggagaccc nnnnnnnnnn nnnnnnnnnn 1020

<210> 21

<211> 1020

<212> DNA

<213> Mus musculus

<220>

<221> Intron

<222> (1)..(584)

<220>

<221> misc_feature

<222> (585)..(620)

<223> exon 5

<220>

<221> Intron

<222> (621)..(1020)

<400> 21

```
tggccacact agctttttac cagttcttcc caggcaaatt ccttagccag gatgtatgtt 60
gctgtatgtt gccttctctg ttacattgta tatttttcat gagccccagc actcgggtgtg 120
taggacttgc ctagcacgtg taagactgat gagagctagt gccctaaagt agttgtagct 180
ggcctagcct tctgggttaa gcaacaccca tgggggctgc tcagaagaag ggatctgagc 240
tgaatgtggc ggctatttcc tgtggggaag aatcctcagc ctgagggtggc tggccgtggc 300
gcttccacct tccccgcctt cctcattgcc cagcttctgg gactgtgggtg gaagaggacc 360
ttcctgtcat gtaacaaaca gctgggtgac tttaaagag agaaagaggg aaaaaaatcc 420
cccaaataaa aacaagaatt gagagtgttt ggggtgccac ttctgttctt cagtgtatgt 480
tgtgggaatc ccctgagaac ccaaacgctt aaggaaaacc actgcagtga agcctttctg 540
agaattaaaa gtatatgacg tttctatttc ttatttgtcc ttagcccaga cggtcttttc 600
ttgtggatat atcacacaaa gtaagtgttc tgaattctgt gtatctattg gatgtctgga 660
```


tcacttgatt tttttttttt agcccctaaa gttgatttcc tctcttcaag ccagccaatg 720
tagtgctcgg gccacagtaa agggaggaga gaggccagga caggaggag gattgctagg 780
gccctggggt cagggtgca actctgctag tccccaaact ggtctttgta gaatagtgat 840
gagttttgct ctcggttctg ctcaggggac tctcctcaaa tattgtcatg gggaccattt 900
ttggttgacg tagggaaaga gcccaggga ctgcatgctg tagtgtgtac ctcagtgt 960
gctgtgaggc actgaggag gacttacgtt cagttccagt nnnnnnnnnn nnnnnnnnnn 1020

<210> 22

<211> 1020

<212> DNA

<213> Mus musculus

<220>

<221> Intron

<222> (1)..(595)

<220>

<221> misc_feature

<222> (596)..(661)

<223> exon 6

<220>

<221> Intron

<222> (662)..(1020)

<400> 22

cagatttcct gagctttcat tgattgggca atgggatttt tttctcagat taatctctat 60
aatacatgca tgtatacaga cacacacaga cacacacgca tgcagtcatt ctcggaagg 120

tgctttttct tattttaata ttaccctctg ttacagccgc tttatgttca ccaggctctt 180
gcatactctgc tgtctcattg gtcattacag atccctctgt ggacggatta ttattgatta 240
ccctcttcag agaagaacgt ggcagttag acagtgtgag tgtatgcaa agtcactcca 300
ctagcaggag gagatcgtga ccacaggctc tcagatctgc agggctcca ccattctgat 360
ttccctgccc cttatccttc aggggtcca gggatgagca gagtgtcag ggctgcccag 420
aaggggcag ctgaggcccc tcaagtccac tctctgcctt tagctcagct gccttttgcg 480
tgtccatggt tcatgagctg catcttgacg ttcacttttt ctagtgtac ccgaccctta 540
aagttcagga ccgcctcgat ttctagatgt gtttatattc ttttcatth cctagaattc 600
cagaggttgg tgagaacaga tggcagaagg actgggtggt cccaagctgg tttaaaaatg 660
ggtaagggat caggatgggt tcctaagtcc ctgaaacca cagaggacc atggcctcct 720
ccctcccttc ttctggctca ctggactcac tcatggagtc tccccattgc tgttgttgtt 780
tttgttgttg ttagcttcta ttgttattgt gaggggtggg gagtgtgttt gtgtgtatga 840
cgtgtgtatg attgcagctg tgtgtacacc atagtactca tcggaggtca gaaggcactt 900
tcaggaggca attctgcctt tcagtagcg gttccagtgt gtgatcacca gactcagatg 960
ctcaggcttt caggacaagc agttttacag gatgagccat nnnnnnnnnn nnnnnnnnnn 1020

<210> 23

<211> 912

<212> DNA

<213> Mus musculus

<220>

<221> Intron

<222> (1)..(389)

<220>

<221> misc_feature

<222> (390)..(491)

<223> exon 7

<220>

<221> Intron

<222> (492)..(912)

<400> 23

ccttcttccc actccctect ccccccttc ctgtctgect tctccttcc tctgcccttt	60
cttctccagt taagggtgaa gttcaggctg aagtggaaat ttcagaatag acacagaaca	120
gaaatgtccc ttggagtact gttctgaaac atctcaccga cttctgaaat aactgagggt	180
tacagggtca ctggaacctc agccccctgac ccacatggtg gccagagagg caaatgctgt	240
accttttatc agaagtgtgt agggatcaag gggtcagtgc cctgagtcct ccagtccacc	300
cagtgggtgtg agtgatgcct tctttccctt gagacacgag tcatggaagc cacctgtcct	360
taccaacttt gtctacctt tgtccacagg acccacagtt atcaagaaga caacgtagaa	420
ggacggagag aaaagggtag aaatggagat cgcattgaag agcctcagct atgggactgg	480
ttcaatccaa agtaaggacg gacaggagat tggggtgggg ggtgctgagt ggggttctga	540
ggagatgctg aggggagtgc tgagggggtg ccggcaggag ggggtgctgg caggagaggg	600
tgctggcagg agggggtgct ggcaggaggg gatgctggca ggagggggtg ctggcaggag	660
gggatgctgg caggaggggg tgctggcagg aggggatgct ggcaggagtg ggtagacctt	720
cctcaatggg ctttggctaa gaaactaaga tctgggtgct ttgaaccaga ctgaacactg	780

tggttaattgc agcaggaaat ggccagtggg aggttaaaca taaacactgg gtgttaagga 840

ctttacaggc cacataggat gctgctgaga aaatgacaag gtctaagggt gagccaagaa 900

nnnnnnnnnn nn 912

<210> 24

<211> 608

<212> DNA

<213> Mus musculus

<220>

<221> Intron

<222> (1)..(221)

<220>

<221> misc_feature

<222> (222)..(359)

<223> exon 8

<220>

<221> Intron

<222> (360)..(608)

<400> 24

catccaggac ctactatctt tgtacttcac tctgtgtcaa gagttggagg taccctacgc 60

atttgtgcct ggcccttgcc aagactcac cccttctgta cttcctgtct ttcatgcagg 120

caagattcag tgacagtcac tggcctccct tccttgacca gtctctcacc acacctcagt 180

gtaatgcttc tgactcgggtg ttgcatgctt cttctcacca ggaaccgccc ggatgttttg 240

acagtgaccc cgtggaaggc gccgattgtg tgggaaggca cttatgacac agctctgctg 300

46

gaaaagtact acgccacaca gaaactcact gtggggctga cagtgtttgc tgtgggaaag 360
taagcaccac tgacaaactc acccttgatg atttgttctt gttctagcat caaaggattt 420
gtgtggggct ccaggggccc acaaaggctg gaatttgaca gtagacttcc cccttctttc 480
ttataatggc tgagaaaaaa caatgatagt aggtgatgag gtatttctct gccagtgagt 540
gagccaatcc aagccagagt agattgtatt aaatacaggt ttattgggaa gctgctctca 600
nnnnnnnn 608

<210> 25
<211> 3240
<212> DNA
<213> Mus musculus

<220>
<221> Intron
<222> (1)..(369)

<220>
<221> misc_feature
<222> (370)..(3010)
<223> exon 9

<220>
<221> terminator
<222> (1088)..(1090)

<220>
<221> 3'UTR
<222> (1091)..(3010)

<220>
<221> Intron

<222> (3011)..(3240)

<400> 25

ttagcagata cactggcctc ttctggatat tcaagagcta gtccttctc tgacagccag 60

cttctcaatc agagaacaga gccttagcat gaaccttact gcaacgcaga gtagttgaga 120

acaccgagct ctcaagtgtg caggcatcga agagcacgcg gtcggggctg tgcacccca 180

gtttgcttaa caaagctggc agtgagataa gtcatgccac tttccccaag gacacaatga 240

ccagctagtg tcgagtggta tgtggagaag ccacccctc ctaacataca atacagatca 300

tctactgtaa tgtaagtat ggtattacat gtatatatgt acccatatat aagtgtgata 360

gtccgtggtg gttcaatgta gccctctcta tttcaggtac attgagcatt acttagaaga 420

ctttctggag tctgtgaca tgtacttcat ggttgccat cgggtcatat tttacgtcat 480

gatagatgac acctcccgga tgctgtcgt gcacctgaac cctctacatt ccttacaagt 540

ctttgagatc aggtctgaga agaggtggca ggatatcagc atgatgcga tgaagaccat 600

tggggagcac atcctggccc acatccagca cgaggtcgac ttcctcttct gcatggacgt 660

ggatcaagtc tttcaagaca acttcggggt ggaaactctg ggccagctgg tagcacagct 720

ccaggcctgg tggtaacaagg ccagtcccga gaagttcacc tatgagaggc gggaactgtc 780

ggccgcgtac attccattcg gagaggggga tttttactac cacgcggcca tttttggagg 840

aacgcctact cacattctca acctcaccag ggagtgcttt aaggggatcc tccaggacaa 900

gaaacatgac atagaagccc agtggcatga tgagagccac ctcaacaaat acttcctttt 960

caacaaaccc actaaaatcc tatctccaga gtattgctgg gactatcaga taggcctgcc 1020

ttcagatatt aaaagtgtca aggtagcttg gcagacaaaa gagtataatt tggttagaaa 1080

taatgtctga cttcaaattg tgatggaaac ttgacactat tactctggct aattcctcaa 1140
acaagtagca acacttgatt tcaactttta aaagaaacaa tcaaaaccaa aaccactac 1200
catggcaaac agatgatttc tcctgacacc ttgagcctgt aatatgtgag aaagagtcta 1260
tggcaagtaa tcaggtataa attctcaatg atttcttata tattctgggt cttgggaaaa 1320
cttgattcta gaaatcaaaa ttaatttgac aaaggaaaag cagatgccgg aaacttcttc 1380
ccagtctgtc atacaattca ccactggcca ggtgctgaga gaagcattag ggaacagtgt 1440
gggttggtgc agagttggac ggctccatcc ctttggttc attatcttcc tctcatgga 1500
gattctaaag caaccagag aggctttgca gccagagacc ttaataagg atgccaatgt 1560
gaccatcagt ctgtaaaagc tgatggctcc aggagcgtg gcagtcagg cccactagg 1620
ctattgtttc tgtcctgggc ataaaggagg cagagagtgc caataggtac tttggtggca 1680
catgttcaga gtccaggaaa aatcaagggt gaccacttag agggacatag gacttggggt 1740
tggtgattga actgagttac aaacacagac agctttcttc aggatgacta acagcaggaa 1800
ttgaatggaa agtgtgttca tttgttttg cccaaattgt attcatgctg ttagctttgt 1860
gtgttgagcc ctgtggagag ggtgtgactg tatcagggaa ggagagtacc tcagcggact 1920
gaggaccagc accctattat atcagaagac aatctctcat catcaggtcc tacctacaac 1980
ctgctctgaa cctccgagtt cctcagccca tcgtgttcca gtgtgggggc ctgtatggag 2040
caggtgactg aagacaaagc cccctgtcac atgacctcat ttccctgct ctagtactat 2100
gcaagtgtga cagccagcca gccagatgta ctggacaaca taggaaccga ctttatggca 2160
atgggagccg cagtcactac aacggagctg ctgaaggttc tggtcccccgc tctgagagcc 2220

tgaggagcc cctgtatagg tggttctcaa cctatgggtc gcgacccctt tgggaagtgt 2280
taaatagccc ttccacaggt gtcccctaag acggttaaaa aacatagata ttccactct 2340
gactggtaac agtagcagaa ttacagttat gaaatagcaa gggaaataat tctggggttc 2400
gtgtcatcca taccatgagg agctacatta ggtcacatca ttagggaagt tgagaagcat 2460
agctctactt gggatattta gcaaattatg caaagggggt tgcgctctg tgttctgtgt 2520
atgcatatat ttatattttg cttgtcttcc agtttaggtc aatctgtttc ttcctttaag 2580
cagtttattt aaaaggccat tgcaccatct tggagaacag catgaggggt ttcaataaaa 2640
aataggatct tacctttgtc cacagggtc tacctcttac tttcaattg tgaacaaaaa 2700
aggctgcaca cccagaggca acaaaaccca cagaattcct gaaccaatgg gagatgccaa 2760
tggaagcaga gcttgacat ctgctaaaaa ttctgcctct ctgtcactgt gctggatccg 2820
tctaaagtgg gacagttcaa tggctctgaaa gtttcaaaaa ggctggggaa tttgagggga 2880
ttttttttta aaataaaatt gatccaagtt taaatctcta atgagtaagc ttaggatttt 2940
attaaaggta atttttagac attcttcaaa ataagaattc ttgtttataa ttgaataaat 3000
tattttctca gtatattttg gtctggtatg gattatgcgt tgtatcctga agatgttcag 3060
aagtgtcagt tgtgattgtc cataatcata aaggatttta cgataccttg aatgagcttc 3120
acaaagacaa gattacaaag aacaggcttt atttcaaat tataaagtgt gctctctctc 3180
aatctctctc tctctctctc nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3240

<210> 26

<211> 3537

<212> DNA

<213> Mus musculus

<400> 26

atcgccccca gaacagatct gactgcctct ttcattcgcc cggaggtaga taggtgtgtc	60
ttaggaggct ggagattctg ggtggagccc tagccctgcc ttttcttagc tggctgacac	120
cttcccttgt agactcttct tggaatgaga agtaccgatt ctgctgaaga cctcgcgctc	180
tcaggctctg ggagttggaa ccctgtacct tcctttcctc tgctgagccc tgccctctta	240
ggcaggccag agctcgacag aagctcgggt gctttgctgt ttgctttgga gggaacacag	300
ctgacgatga ggctgacttt gaactcaaga gatctgctta cccagctctc ctggaattaa	360
aggcctgtac tacatttgcc tggacctaa gattttcatga tcactatgct tcaagatctc	420
catgtcaaca agatctccat gtcaagatcc aagtcagaaa caagtcttcc atcctcaaga	480
tctggatcac aggagaaaat aatgaatgtc aagggaagaa taatcctgtt gatgctgatt	540
gtctcaaccg tggttgtcgt gttttgggaa tatgtcaaca gccagacgg ctctttcttg	600
tggatatatc acacaaaaat tccagagggt ggtgagaaca gatggcagaa ggactggtgg	660
ttcccaagct ggtttaaaaa tgggaccac agttatcaag aagacaacgt agaaggacgg	720
agagaaaagg gtagaatgg agatcgcat gaagagcctc agctatggga ctggttcaat	780
ccaagaacc gcccgatgt ttgacagtg acccgtgga aggcgccgat tgtgtgggaa	840
ggcacttatg acacagctct gctggaaaag tactacgcca cacagaaact cactgtgggg	900
ctgacagtgt ttgctgtgg aaagtacatt gacattact tagaagactt tctggagtct	960
gctgacatgt acttcatggt tggccatcgg gtcattttt acgtcatgat agatgacacc	1020

tcccggatgc ctgtcgtgca cctgaacct ctacattcct tacaagtctt tgagatcagg 1080
tctgagaaga ggtggcagga tatcagcatg atgcgcatga agaccattgg ggagcacatc 1140
ctggcccaca tccagcacga ggtcgacttc ctcttctgca tggacgtgga tcaagtcttt 1200
caagacaact tcggggtgga aactctgggc cagctggtag cacagctcca ggcctggtgg 1260
tacaaggcca gtcccagaa gttcacctat gagaggcggg aactgtcggc cgcgtacatt 1320
ccattcggag agggggattt ttactaccac gcggccattt ttggaggaac gcctactcac 1380
attctcaacc tcaccaggga gtgctttaag gggatcctcc aggacaagaa acatgacata 1440
gaagcccagt ggcgatga gagccacctc acaaatact tccttttcaa caaaccact 1500
aaaatcctat ctccagagta ttgctgggac tatcagatag gctgccttc agatattaaa 1560
agtgtcaagg tagcttgga gacaaaagag tataatttgg ttagaaataa tgtctgactt 1620
caaattgtga tggaaacttg aactattac tctggctaatt tcctcaaaca agtagcaaca 1680
cttgatttca acttttaaaa gaaacaatca aaacaaaac ccactaccat ggcaaacaga 1740
tgatttctcc tgacaccttg agcctgtaat atgtgagaaa gagtctatgg caagtaatca 1800
ggtataaatt ctcaatgatt tcttatatat tctgggtctt gggaaaactt gattctagaa 1860
atcaaaatta atttgacaaa ggaaaagcag atgccggaaa cttcttccca gtctgtcata 1920
caattcacca ctggccaggt gctgagagaa gcattaggga acagtgtggg ttgtgtcaga 1980
gttgacggc tccatccctt tggttcatt atcttctcc tcatggagat tctaaagcaa 2040
cccagagagg ctttgacagc agagacctt aataaggatg ccaatgtgac catcagctcg 2100
taaaagctga tggctccagg agcgtggca gtccaggccc cactaggcta ttgtttctgt 2160

cctgggcata aaggaggcag agagtgccaa taggtacttt ggtggcacat gttcagagtc 2220

caggaaaaat caagggtgac cacttagagg gacataggac ttggggtttg tgattgaact 2280

gagttacaaa cacagacagc tttcttcagg atgactaaca gcaggaattg aatggaaagt 2340

gtgttcattt tgttttgccc aaattgtatt catgctgtta gctttgtgtg ttgagccctg 2400

tggagagggg gtgactgtat cagggaagga gagtacctca gcggactgag gaccagcacc 2460

ctattatata agaagacaat ctctcatcat caggtcctac ctacaacctg ctctgaacct 2520

ccgagttcct cagcccatcg tgttcacgtg tgggggacctg tatggagcag gtgactgaag 2580

acaaagcccc ctgtcacatg acctcatttc ccctgctcta gtactatgca agtgtgacag 2640

ccagccagcc agatgtactg gacaacatag gaaccgactt tatggcaatg ggagccgcag 2700

tactacaac ggagctgctg aaggttctgt tccccgctct gagagcctgc aggagccctt 2760

gtataggtgg ttctcaacct atgggtcgcg accccttttg gaagtgttaa atgacccttt 2820

cacaggtgtc ccctaagacg gttaaaaaac atagatattt ccactctgac tggtaacagt 2880

agcagaatta cagttatgaa atagcaaggg aaataattct ggggttcgtg tcatccatac 2940

catgaggagc tacattaggt cacatcatta gggaagttga gaagcatagc tctacttggg 3000

tatttaagca aattatgcaa aggggggttg cgctctgtgt tctgtgtatg catatattta 3060

tattttgctt gtcttcacgt ttaggtcaat ctgtttcttc cttaagcag tttattttaa 3120

aggccattgc accatcttgg tgaacagcat gaggggtttc aataaaaaat aggatcttac 3180

ctttgtccac agggctctac ctcttacttt tcaattgtga acaaaaaagg tcgcacaccc 3240

agaggcaaca aaaccacag aattcctgaa ccaatgggag atgccaatgg aagcagagct 3300

53

tgacacatctg ctaaaaattc tgcctctctg tcaactgtgct ggatccgtct aaagtgggac 3360
agttcaatgg tctgaaagtt tcaaaaaggc tggggaattt gaggggattt ttttttaaaa 3420
taaaattgat ccaagtttaa atctctaatt agtaagctta ggattttatt aaaggtaatt 3480
tttagacatt cttcaaaata agaattcttg tttataattg aataaattat tttctca 3537

<210> 27

<211> 3135

<212> DNA

<213> Homo sapiens

<400> 27

gtggctgac agagcgcgta gggcttcgcc ggggccggga gctgggcgcg gtcctgctca 60
gccagctca ccgcgcgcg gccctcggcg ccctcggcgc cctggttctg cggatcagga 120
gaaaataatg aatgtcaaag gaaaagtaat tctgtcaatg ctggttgtct caactgtgat 180
cattgtgttt tgggaattta tcaacagcac agaaggctct ttcttggtga tatatcactc 240
aaaaaaccca gaagttgatg acagcagtgc tcagaagggc tggtggtttc tgagctgggt 300
taacaatggg atccacaatt atcaacaagg ggaagaagac atagacaaag aaaaaggaag 360
agaggagacc aaaggaagga aatgacaca acagagcttc ggctatggga ctggtttaat 420
ccaaaatata atgatcatta cttggaggag ttcataacat ctgctaatag gtacttcatg 480
gttgccaca aagtcattt ttacatcatg gtggatgatg tctccaagct gccgtttata 540
gagctgggtc ctctgcatc cttcaaaatg tttgaggtca agccagagaa gaggtggcaa 600
gacatcagca tgatgcgtat gaagatcact ggggagcaca tcttgccca catccaacac 660
gaggtcgact tcctcttctg catggatgtg gaccaggtct tccaagacca tttgggggtg 720

gagaccctag gccagtcagt ggctcagcta caggctggcg gtacaaggca gatccctatg 780

actttaccta ggagaggtgg aaagagtcag caggatacat tccatttggc caggggattt 840

ttattacat gcagccattt ctggaggaac acccattcag gttctcaaca tcaccagga 900

gtgctttaag ggaatcctcc tggacaagaa aaatgacata gaagccaagt ggcattgatga 960

aagccaccta aacaagtatt tccttctcaa taaacctctt aaaatcttat ccctaaaata 1020

ctgctgggat tatcatatag gcctgccttc agatattaaa actgtcaagt gatcgtggca 1080

gacaaaagag tataatttgg ttagaaataa tgtctgactt caaattgtgc cagtagattt 1140

ctgaatttaa gagagagaat attctggcta cttcctcaga aaagtaacac ttaattttaa 1200

cttcaaaaaa tactaatgaa acaccaacag ggcaaaaaca taccattcct ccttgtaact 1260

tggggctttg taatgtggaa gaatgaatct agggcaatca gatataaatt ccagtgatt 1320

tcttatctat tctgggtttg ggggaaatac tatcaactga accaaaaata acttgtcata 1380

ggcagagata aagccagaaa cactctacac atgccagatg acatctggag aaaaggtgc 1440

taaggaagc gtttggcagc aagatatgat tgtaaggggt tgtcccttga gttcaatgct 1500

tgctatttc tgatgggtct aaagcaacat ggagttactg tgcagcagaa ctctcagtaa 1560

agacaccatt tgccttgga atcctcaaaa agcttcaata gcagattgct tcagaccatc 1620

tgtagtccgt ccttttctca tctggatgtt gtttggcttc tgtgcgaaag attggtggag 1680

tgtcccagta gatatcatgg tgggtgtgtga tcagagtccc aaggaacctg aatgagccaa 1740

ggtgcccagc atgaagtcaa aacaaagcct tgacatgagt ttgccatgaa atagcgaaga 1800

gagagtggaa gagaggagcc aatcactgtg gggcagtgcc accctgaggg cacttagggt 1860

atgggggttg tgcttaaata catcacagat ccaggtactg aatgggagga agtgtgggtg 1920

atttccaatc tcattgacct tatgttcagg gacttgaacg gaagatgttt cttgtgttgc 1980

ctaagtggta ttcagtctac cagactctgc aacttgcac ttcaaatcct tggtaaagag 2040

atgtggatgg tgcagagaa ggcaaaggcc tgcagtggat tgaagaggct tgcaagcagt 2100

tctgtttcta ggatgtgggc ttcacagaa gacactcggc caccacttag ctagtctaaa 2160

cctcagggtt cctcagccca tcatacccca acttggagga ctgacatcaa ggagtagact 2220

ggagaaacag ccctcccatc aagtaacctc ttgttctctc ctgctccatc tgcactatag 2280

aagtgttaata attagacata cttggcaaaa tggctaattg atttggtaac agaagcatga 2340

gccataacaa tggaagatct agttatcatg actgaacagc ttaacattca attcccttct 2400

ctaagagaag ctgtgaaatc ctacatatta tttaaagtta accaaatcaa tgtaaaggga 2460

gttaggagac agtgtgtacc tatgcacgta tatttatgtt ttgcttgtgt tccagtctcg 2520

gtcatttgtt tccattttca agcaatttat ttgaagagcc attgcactag cctgatgtat 2580

actgcaatga gcttctttga taaaatgaaa cttaaatttt tctcgaccat ttcaccgtgc 2640

ctoctacttc attttttgcc agaaaatctc acatccaaca aaacaaaaca aaaaccctga 2700

attagtgggc ttgaaaagg aaaaagcagg gctttgaaaa agtagatcac acatcagtta 2760

agactcctgc ttctctatta gtcaggttgt cttggattca gtctggagta ggcagagctt 2820

aagggttttt aagtcctgac ccaaagaaat gatctagcct gaaagttag agcaaaggac 2880

taatgtttac ttttaaagga atttcttgat ttttttaaaa aacttcatta aagtttaaat 2940

ccccaatgga caaattcata atcttgtaa tcgttattac taaacttttt aaaaaatgtc 3000

ccaatttaca attaaataaa ttactttctc agtatattct ggtctggtca tggattgtgc 3060
atttcctccc aaagatattc aaaattgtca attagagaat tttaggtttt cagactcaga 3120
aaagtctctca cgccc 3135

<210> 28

<211> 3558

<212> DNA

<213> Homo sapiens

<400> 28

ctcagactga atacatggcc cactgtcgct ccagccatct caaatggaac gacctgttct 60
ctgaagtata tcttacagtg ctttctctcg aatccccttt gggaaatcta aaggctgaat 120
ccagccagct ttccatgct gcctggtctg gaaatcactg caagggtttt tcccagagaa 180
ccaaagtaag ataaatgaaa gatgctacac aattctgctg agggctctgt ctactcccca 240
tctcctgaaa cagctgttta ttctttcgac aggagttgaa accagcacct tccctttctc 300
tgagtctctgc ctccttctgc ggaaggaggc tcaaaagaac tttgttggtt tgccttttac 360
tctggggtga aagcggcagg aggtatgtga gatggtgaaa tgatttgctt ctgccatgct 420
ggggtcacgg gtggatcgcc ctaaactctc ggtggccccc tcagtagttt tggaagagga 480
ccaagtcctt gtctctccag cagtggacct ggaagaagga tgccggctca gggacttcac 540
tgagaaaata atgaatgtca aaggaaaagt aattctgtca atgctggttg tctcaactgt 600
gatcattgtg ttttggaat ttatcaacag cacagaaggc tctttcttgt ggatatatca 660
ctcaaaaaac ccagaagttg atgacagcag tgctcagaag ggctggtggt ttctgagctg 720

gtttaacaat gggatccaca attatcaaca aggggaagaa gacatagaca aagaaaaagg 780

aagagaggag accaaaggaa ggaaaatgac acaacagagc ttcggctatg ggactggttt 840

aatccaaaat ataatgatca ttacttggag gagttcataa catctgctaa taggtacttc 900

atggttggcc acaaagtcac attttacatc atggtggatg atgtctccaa gctgccgttt 960

atagagctgg gtccctctgca ttccctcaaa atgtttgagg tcaagccaga gaagagggtg 1020

caagacatca gcatgatgag tatgaagatc actggggagc acatcttggc ccacatccaa 1080

cacgaggctg acttcctctt ctgcatggat gtggaccagg tcttccaaga ccattttggg 1140

gtggagaccc taggccagtc agtggctcag ctacaggctg gcggtacaag gcagatccct 1200

atgactttac ctaggagagg tggaagagt cagcaggata cattccattt ggccagggga 1260

tttttattac catgcagcca tttctggagg aacacccatt caggttctca acatcaccca 1320

ggagtgcctt aagggaatcc tctggacaa gaaaaatgac atagaagcca agtggcatga 1380

tgaaagccac ctaaacaagt atttccttct caataaaccc tctaaaatct tatccctaaa 1440

atactgctgg gattatcata taggcctgcc ttcagatatt aaaactgtca agtgatcgtg 1500

gcagacaaaa gagtataatt tggtagaaa taatgtctga cttcaaattg tgccagtaga 1560

tttctgaatt taagagagag aatattctgg ctacttctc agaaaagtaa cacttaattt 1620

taacttcaaa aaatactaata gaaacaccaa cagggcaaaa acataccatt cctccttgta 1680

acttggggct ttgtaatgtg gaagaatgaa tctagggcaa tcagatataa attcccagtg 1740

atttcttacc tattctgggt ttgggggaaa tactatcaac tgaacaaaaa ataacttgtc 1800

ataggcagag ataaagccag aaacactcta cacatgccag atgacatctg gagaaaaggg 1860

tgctaaggga agcgtttggc agcaagatat gattgtaagg ggttgccct tgagttcaat 1920
gtctgcctat ttctgatggg tctaaagcaa catggagtta ctgtgcagca gaactctcag 1980
taaagacacc atttgccttg gcaatcctca aaaagcttca atagcagatt gcttcagacc 2040
atctgtagtc cgtccttttc tcatctggat gttgtttggc ttctgtgcga aagattggtg 2100
gagtgtccca gtagatatca tgggtggtgtg tgatcagagt cccaaggaac ctgaatgagc 2160
caagggtgcc agcatgaagt caaaacaaag ccttgacatg agtttgccat gaaatagcga 2220
agagagagtg gaagagagga gccaatcact gtggggcagt gccaccctga gggcacttag 2280
ggtatggggt tgggtgcttaa atacatcaca gatccaggta ctgaatggga ggaagtgtg 2340
gtgatttcca atctcattga ccctatgttc agggaactga acggaagatg tttcttgtgt 2400
tgcctaagtg gtattcagtc taccagactc tgcaacttgc atcttcaa at ccttggtaaa 2460
gagatgtgga tgggtgtcaga gaaggcaaag gcctgcagtg gattgaagag gcttgcaagc 2520
agttctgttt ctaggatgtg ggcttcatca gaagacactc ggtcaccact tagctagtct 2580
aaacctcagg gttcctcagc ccatacacc ccaacttga ggactgacat caaggagtag 2640
actggagaaa cagccctccc atcaagtaac ctctgttct ctctgctcc atctgacta 2700
tagaagtgtg ataattagac atacttggca aaatggctaa ttgatttggg aacagaagca 2760
tgagccataa caatggaaga tctagttatc atgactgaac agcttaacat tcaattccct 2820
tctctaagag aagctgtgaa atcctacata ttatttaaag ttaaccaa at caatgtaaag 2880
ggagttagga gacagtgtgt acctatgcac gtatatttat gttttgcttg tgttcagtc 2940
tcggtcattt gtttccattt tcaagcaatt tatttgaaga gccattgcac tagcctgatg 3000

tatactgcaa tgagcttctt tgataaaatg aaacttaaata ttttctcgac catttcaccg 3060
tgctcctac ttcatttttt gccagaaaat ctcacatcca acaaaacaaa acaaaaaccc 3120
tgaattagtg ggctttgaaa aggaaaaagc agggctttga aaaagtagat cacacatcag 3180
ttaagactcc tgcttctcta ttagtcaggt tgtcttggat tcagtctgga gtaggcagag 3240
cttaagggtt ttttaagcct gacccaaaga aatgatctag cctgaaagtt tagagcaaag 3300
gactaatggt tactttttaa ggaatttctt gattttttta aaaaacttca ttaaagttta 3360
aatccccaat ggacaaattc ataatcttgt taatcgttat tactaaactt tttaaaaaat 3420
gtcccaattt acaattaaat aaattacttt ctcagtatat tctggtctgg tcatggattg 3480
tgcatctcct ccaaagata ttcaaaattg tcaattagag aattttaggt tttcagactc 3540
agaaaagtcc tcacgcc 3558

<210> 29

<211> 852

<212> DNA

<213> Homo sapiens

<400> 29

gtggctgac agagcgcgta gggcttcgcc ggggcccggga gctgggcccgt gtcctgctca 60
gccagctca ccgcgcgcgc gccctcggcg cctcggcgc cctggttctg cggatcagga 120
gaaaataatg aatgtcaaag gaaaagtaat tctgtcaatg ctggttctct caactgtgat 180
cattgtgttt tgggaattta tcaacagcac agaaggctct ttcttgtgga tatatcactc 240
aaaaaaccca gaagttgatg acagcagtgc tcagaagggc tgggtggttc tgagctggtt 300
taacaatggg atccacaatt atcaacaagg ggaagaagac atagacaaag aaaaaggaag 360

agaggagacc aaaggaagga aaatgacaca acagagcttc ggctatggga ctggtttaat 420
ccaaaacttga aggaatccga ataactaaac tggactctgg ttttctgact cagtccttct 480
agaagacctg gactgagaga tcatgcggtt aaggagtgtg taacaggcgg accacctggt 540
gggactgcga gattctcaag gggaaggact gggctctcatt tctcccatct cagcgcttag 600
caggatgacc tggtatagag cagggaactg ggaaatgtgg gtcaggggat cagacactcc 660
agttgggtct tttatataaa ttaaattggca aaaggctcca tacccttctc cttctttcct 720
accctccact ttatctgcaa aatgggaatg atgataacac ccacttcata gaatggtcac 780
gaagatcaaa tgagagaata aaagtcaagc acttagcctc tggcgacaaa taagtattaa 840
ataagtatac ct 852

<210> 30

<211> 1232

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)..(118)

<223> This is exon 1

<220>

<221> misc_feature

<222> (119)..(207)

<223> This is exon 4

<220>

<221> .misc_feature
<222> (208)..(243)
<223> This is exon 5

<220>
<221> misc_feature
<222> (244)..(309)
<223> This is exon 6

<220>
<221> misc_feature
<222> (310)..(425)
<223> This is exon 7

<220>
<221> misc_feature
<222> (426)..(1232)
<223> This is exon 8b

<220>
<221> misc_feature
<222> (797)..(802)
<223> putative polyadenylation signal

<400> 30
gtggctgatc agagcgcgta gggcttcgcc ggggccggga gctgggcgcg gtcctgctca 60

gccagctca ccgcgcgcgc gccctcggcg ccctcggcgc cctggttctg cggatcagga 120

gaaaataatg aatgtcaaag gaaaagtaat tctgtcaatg ctggttgtct caactgtgat 180

cattgtgttt tgggaattta tcaacagcac agaaggctct ttcttgtgga tatatcactc 240

aaaaaaccca gaagttgatg acagcagtgc tcagaagggc tggtaggttc tgagctggtt 300

taacaatggg atccacaatt atcaacaagg ggaagaagac atagacaaag aaaaaggaag 360

agaggagacc aaaggaagga aaatgacaca acagagcttc ggctatggga ctggtttaat 420

ccaaacttga aggaatccga ataactaaac tggactctgg tttctgact cagtccttct 480

agaagacctg gactgagaga tcatgcggtt aaggagtgtg taacaggcgg accacctgtt 540

gggactgcga gattctcaag gggaaggact gggctctcatt tctcccatct cagcgcttag 600

caggatgacc tggtagagag cagggaactg ggaaatgtgg gtcaggggat cagacactcc 660

agttgggtct tttatataaa ttaaattggca aaaggctcca tacccttctc cttctttcct 720

accctccact ttatctgcaa aatgggaatg atgataacac ccacttcata gaatggtcat 780

gaagatcaaa tgagagaata aaagtcaagc acttagcctc tggtagcaca taagtattaa 840

ataagtatac ctattcctcc ttttcctttt ttaaaaataa tattaccaa tgtccagctt 900

atacacattt acaagactta gctagtgggc tatgttagag ctactaaaag atctttgaca 960

agctaaaact aagatgcaat gaatgaggtg taacgaacaa gagagtttta agttcagaaa 1020

tggttacaga agtataagac agctgtgtgg gtgttttttg gtttttggtt tctggtttac 1080

aatctcgtca ttcaacaaag atgggagttt tatagaacta aaagcaccat gtaagctact 1140

aaaaacaaca acaaaaaagg ctcatcattt ctgagtctga attgacaaaa atgccaatgc 1200

aaataaaaat gattactttt tatttttcaa cg 1232

<210> 31

<211> 1275

<212> DNA

<213> Homo sapiens

<400> 31

ctcagactga atacatggcc cactgtcgct ccagccatct caaatggaac gacctgttct	60
ctgaagtata tcttacagtg ctttctctcg aatccccctt gggaaatcta aaggctgaat	120
ccagccagct tttccatgct gcctggctcg gaaatcactg caagggtttt tcccagagaa	180
ccaaagtaag ataaatgaaa gatgctacac aattctgctg agggctctgt ctactcccca	240
tctcctgaaa cagctgttta ttctttcgac aggagttgaa accagcacct tccctttctc	300
tgagtctgc ctcttctgc ggaagggagc taaaagaac ttgtttgtt tgccttttac	360
tctggggtga aagcggcagg aggtatgtga gatggtgaaa tgatttgctt ctgccatgct	420
ggggtcacgg gtggatcgcc ctaaactctc ggtggccccc tcagtagttt tggaagagga	480
ccaagtcctt gtctctccag cagtggacct ggaagaagga tgccggctca ggacttcac	540
tgagaaaata atgaatgtca aaggaaaagt aattctgtca atgctggtt tctcaactgt	600
gatcattgtg ttttgggaat ttatcaacag cacagaaggc tctttcttgt ggatatatca	660
ctcaaaaaac ccagaagttg atgacagcag tgctcagaag ggctgggtgt ttctgagctg	720
gtttaacaat gggatccaca attatcaaca aggggaagaa gacatagaca aagaaaaagg	780
aagagaggag accaaaggaa ggaaaatgac acaacagagc ttcggctatg ggactggttt	840
aatccaaact tgaaggaatc cgaataacta aactggactc tggttttctg actcagtcct	900
tctagaagac ctggactgag agatcatgcg gttaaggagt gtgtaacagg cggaccacct	960
gttgggactg cgagattctc aagggaagg actgggtctc atttctcca tctcagcgct	1020
tagcaggatg acctggtata gagcagggaa ctgggaaatg tgggtcaggg gatcagacac	1080

tccagttggg tcttttatat aaattaaatg gcaaaaggct ccataccctt ctccttcttt 1140
cctaccctcc actttatctg caaaatggga atgatgataa caccacttc atagaatggt 1200
catgaagatc aaatgagaga ataaaagtca agcacttagc ctctggtgca caataagtat 1260
taaataagta tacct 1275

<210> 32
<211> 1655
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(272)
<223> This is exon 1a

<220>
<221> misc_feature
<222> (273)..(541)
<223> This is exon 2

<220>
<221> misc_feature
<222> (542)..(630)
<223> This is exon 4

<220>
<221> misc_feature
<222> (631)..(666)
<223> This is exon 5

<220>

<221> misc_feature

<222> (667)..(732)

<223> This is exon 6

<220>

<221> misc_feature

<222> (733)..(848)

<223> This is exon 7

<220>

<221> misc_feature

<222> (849)..(1655)

<223> This is exon 8h

<400> 32

ctcagactga atacatggcc cactgtcgct ccagccatct caaatggaac gacctgttct 60

ctgaagtata tcttacagtg ctttctctcg aatccccctt gggaatcta aaggctgaat 120

ccagccagct tttccatgct gcttggctcg gaaatcactg caagggtttt tcccagagaa 180

ccaaagtaag ataaatgaaa gatgctacac aattctgctg agggctctgt ctactcccca 240

tctctgaaa cagctgttta ttctttcgac aggagttgaa accagcacct tccctttctc 300

tgagtctgc ctccttctgc ggaaggagc tcaaaagaac ttgttggtt tgccttttac 360

tctggggtga aagcggcagg aggtatgtga gatggtgaaa tgatttgctt ctgccatgct 420

ggggtcacgg gtggatcgcc ctaaactctc ggtggccccc tcagtagttt tggaagagga 480

ccaagtcctt gtctctccag cagtggacct ggaagaagga tgccggctca gggacttcac 540

tgagaaaata atgaatgtca aaggaaaagt aattctgtca atgctggttg tctcaactgt 600
gatcattgtg ttttggaat ttatcaacag cacagaaggc tctttcttgt ggatatatca 660
ctcaaaaaac ccagaagttg atgacagcag tgctcagaag ggctggtggt ttctgagctg 720
gtttaacaat gggatccaca attatcaaca aggggaagaa gacatagaca aagaaaaagg 780
aagagaggag accaaaggaa ggaaaatgac acaacagagc ttcggctatg ggactggttt 840
aatccaaact tgaaggaatc cgaataacta aactggactc tggttttctg actcagtcct 900
tctagaagac ctggactgag agatcatgcg gttaaggagt gtgtaacagg cggaccacct 960
gttgggactg cgagattctc aaggggaagg actgggtctc atttctcca tctcagcgct 1020
tagcaggatg acctggtata gacagggaa ctgggaaatg tgggtcaggg gatcagacac 1080
tccagttggg tcttttatat aaattaaatg gcaaaaggct ccataccctt ctcttcttt 1140
cctaccctcc actttatctg caaaatggga atgatgataa caccacttc atagaatggt 1200
catgaagatc aaatgagaga ataaaagtca agcacttagc ctctggtgca caataagtat 1260
taaataagta tacctattcc tccttttctt tttttaaaaa taatattacc aaatgtccag 1320
cttatacaca ttacaagac ttagctagtg ggctatgtta gagctactaa aagatctttg 1380
acaagctaaa actaagatgc aatgaatgag gtgtaacgaa caagagagtt ttaagttcag 1440
aaatggttac agaagtataa gacagctgtg tgggtgtttt ttggtttttg gtttctggtt 1500
tacaatctcg tcattcaaca aagatgggag ttttatagaa ctaaaagcac catgtaagct 1560
actaaaaaca acaacaaaaa aggtcatca tttctcagtc tgaattgaca aaaatgccaa 1620
tgcaataaaa atgattact ttttattttt caacg 1655

<210> 33

<211> 3322

<212> DNA

<213> Homo sapiens

<400> 33

gtggctgac agagcgcgta gggcttcgcc ggggccggga gctgggcgcg gtcctgctca	60
gccagctca ccgcgcgcg gccctcggcg cctcgggcg cctggttctg cggatcagga	120
gaaaataatg aatgtcaaag gaaaagtaat tctgtcaatg ctggttgtct caactgtgat	180
cattgtgttt tgggaattta tcaacagcac agaaggctct ttcttggtga tatatcactc	240
aaaaaaccca gaagttgatg acagcagtgc tcagaagggc tgggtggttc tgagctggtt	300
taacaatggg atccacaatt atcaacaagg ggaagaagac atagacaaag aaaaaggaag	360
agaggagacc aaaggaagga aaatgacaca acagagcttc ggctatggga ctggtttaat	420
ccaaacttga aggaatccga ataactaac tggactctgg ttttctgact cagtccttct	480
agaagacctg gactgagaga tcatgcggtt aaggagtgtg taacaggcgg accacctggt	540
gggactgcga gattctcaag gggaaggact gggcttcatt tctcccatct cagcgcttag	600
caggatgacc tgatataatg atcattactt ggaggagttc ataacatctg ctaataggta	660
cttcatggtt ggccacaaag tcatatttta catcatggtg gatgatgtct ccaagctgcc	720
gtttatagag ctgggtcctc tgcattcctt caaatgttt gaggtcaagc cagagaagag	780
gtggcaagac atcagcatga tgcgtatgaa gatcactggg gagcacatct tggccacat	840
ccaacacgag gtcgacttcc tcttctgcat ggatgtggac caggctcttc aagaccattt	900
tggggtggag accctaggcc agtcagtggc tcagctacag gctggcggta caaggcagat	960

ccctatgact ttacctagga gaggtggaaa gagtccagcag gatacattcc atttggccag 1020

gggattttta ttaccatgca gccattttctg gaggaacacc cattcaggtt ctcaacatca 1080

cccaggagtg ctttaaggga atcctcctgg acaagaaaaa tgacatagaa gccaaagtggc 1140

atgatgaaag ccacctaaac aagtatttcc ttctcaataa accctctaaa atcttatccc 1200

taaaatactg ctgggattat catataggcc tgccttcaga tattaaaact gtcaagtgat 1260

cgtggcagac aaaagagtat aatttggtta gaaataatgt ctgacttcaa attgtgccag 1320

tagattttctg aatttaagag agagaatatt ctggctactt cctcagaaaa gtaacactta 1380

attttaactt caaaaaatac taatgaaaca ccaacagggc aaaaacatc cattcctcct 1440

tgtaacttgg ggctttgtaa tgtggaagaa tgaatctagg gcaatcagat ataaattccc 1500

agtgatttct tatctattct gggtttgggg gaaatactat caactgaacc aaaaataact 1560

tgtcataggc agagataaag ccagaaacac tctacacatg ccagatgaca tctggagaaa 1620

agggtgctaa gggaagcgtt tggcagcaag atatgattgt aaggggttgt cccttgagtt 1680

caatgtctgc ctatttctga tgggtctaaa gcaacatgga gttactgtgc agcagaactc 1740

tcagtaaaga caccatttgc cttggcaatc ctcaaaaagc ttcaatagca gattgcttca 1800

gaccatctgt agtcctcct tttctcatct ggatgttgtt tggcttctgt gcgaaagatt 1860

ggtggagtgt ccagtagat atcatggtgg tgtgtgatca ggtcccaag gaacctgaat 1920

gagccaaggt gccagcatg aagtcaaaac aaagccttga catgagtttg ccatgaaata 1980

gcgaagagag agtggaagag aggagccaat cactgtgggg cagtgccacc ctgagggcac 2040

ttagggatg gggttggtgc ttaaatacat cacagatcca ggtactgaat gggaggaagt 2100

gtgggtgatt tccaatctca ttgaccctat gttcaggac ttgaacggaa gatgtttctt 2160

gtgttgcta agtgggtatc agtctaccag actctgcaac ttgcatcttc aaatccttgg 2220

taaagagatg tggatggtgt cagagaaggc aaaggcctgc agtggattga agaggcttgc 2280

aagcagttct gtttctagga tgtgggcttc atcagaagac actcggtcac cacttagcta 2340

gtctaaacct cagggttctc cagcccatca taccccaact tggaggactg acatcaagga 2400

gtagactgga gaaacagccc tcccatcaag taacctcttg ttctctctg ctccatctgc 2460

actatagaag tgtaataatt agacatactt ggcaaatgg ctaattgatt tggtaacaga 2520

agcatgagcc ataacaatgg aagatctagt tatcatgact gaacagctta acattcaatt 2580

cccttctcta agagaagctg tgaaatccta catattattt aaagttaacc aaatcaatgt 2640

aaaggagtt aggagacagt gtgtacctat gcaagtatat ttatgttttg cttgtgttcc 2700

agtctcggtc atttgtttcc attttcaagc aattttattg aagagccatt gcactagcct 2760

gatgtatact gcaatgagct tctttgataa aatgaaactt aaatttttct cgaccatttc 2820

accgtgcctc ctacttcatt ttttgccaga aaatctcaca tccaacaaaa caaaacaaaa 2880

accctgaatt agtgggcttt gaaaaggaaa aagcagggtt ttgaaaaagt agatcacaca 2940

tcagttaaga ctctgcttc tctattagtc aggttgtctt ggattcagtc tggagtaggc 3000

agagcttaag ggtttttaag tcctgaccca aagaaatgat ctacgctgaa agtttagagc 3060

aaaggactaa tgtttacttt taaaggaatt tcttgatttt tttaaaaaac ttcattaaag 3120

tttaaatccc caatggacaa attcataatc ttgttaatcg ttattactaa actttttaaa 3180

aatgtccca atttacaatt aaataaatta ctttctcagt atattctggt ctgggtcatgg 3240

attgtgcatt tcctcccaa gatattcaaa attgtcaatt agagaatttt aggttttcag 3300

actcagaaaa gtcctcacgc cc 3322

<210> 34

<211> 3745

<212> DNA

<213> Homo sapiens

<400> 34

ctcagactga atacatggcc cactgtcgct ccagccatct caaatggaac gacctgttct 60

ctgaagtata tcttacagt cttctctctg aatccccctt gggaaatcta aaggctgaat 120

ccagccagct tttccatgct gcctggctctg gaaatcactg caagggtttt tccagagaa 180

ccaaagtaag ataaatgaaa gatgctacac aattctgtctg agggctctgt ctactcccca 240

tctcctgaaa cagctgttta ttctttcgac aggagttgaa accagcacct tccctttctc 300

tgagtctctg ctccttctgc ggaagggagc tcaaaagaac ttgttggtt tgccttttac 360

tctggggtga aagcggcagg aggtatgtga gatggtgaaa tgatttgctt ctgccatgct 420

ggggtcacgg gtggatcgcc ctaaactctc ggtggccccc tcagtagttt tggaagagga 480

ccaagtcctt gtctctccag cagtggacct ggaagaagga tgccggctca gggacttcac 540

tgagaaaata atgaatgtca aaggaaaagt aattctgtca atgctggttg tctcaactgt 600

gatcattgtg ttttggaat ttatcaacag cacagaaggc tctttcttgt ggatatatca 660

ctcaaaaaac ccagaagttg atgacagcag tgctcagaag ggctgggtgt ttctgagctg 720

gtttaacaat gggatccaca attatcaaca aggggaagaa gacatagaca aagaaaaagg 780

aagagaggag accaaaggaa ggaaaatgac acaacagagc ttcggctatg ggactggttt 840

aatccaaact tgaaggaatc cgaataacta aactggactc tggttttctg actcagtcct 900

tctagaagac ctggactgag agatcatgcg gttaaggagt gtgtaacagg cggaccacct 960

gttgggactg cgagattctc aaggggaagg actgggtctc atttctcca tctcagcgct 1020

tagcaggatg acctgatata atgatcatta cttggaggag ttcataacat ctgctaatag 1080

gtacttcatg gttggccaca aagtcattt ttacatcatg gtggatgatg tctccaagct 1140

gccgtttata gagctgggtc ctctgcattc cttcaaaatg tttgagggtca agccagagaa 1200

gaggtggcaa gacatcagca tgatgcgtat gaagatcact ggggagcaca tcttggccca 1260

catccaacac gaggtcgact tcctcttctg catggatgtg gaccaggctt tccaagacca 1320

ttttggggtg gagaccctag gccagtcagt ggctcagcta caggctggcg gtacaaggca 1380

gatccctatg actttaccta ggagaggtgg aaagagtcag caggatacat tccatttggc 1440

caggggattt ttattaccat gcagccattt ctggaggaac acccattcag gttctcaaca 1500

tcaccagga gtgctttaag ggaatcctcc tggacaagaa aaatgacata gaagccaagt 1560

ggcatgatga aagccaccta aacaagtatt tccttctcaa taaaccctct aaaatcttat 1620

ccctaaaata ctgctgggat tatcatatag gcctgccttc agatattaaa actgtcaagt 1680

gatcgtggca gacaaaagag tataatttgg ttagaaataa tgtctgactt caaattgtgc 1740

cagtagattt ctgaatttaa gagagagaat attctggcta cttcctcaga aaagtaacac 1800

ttaattttaa cttcaaaaaa tactaatgaa acaccaacag ggcaaaaaca taccattcct 1860

ccttgtaact tggggctttg taatgtggaa gaatgaatct agggcaatca gatataaatt 1920

cccagtgatt tcttatctat tctgggtttg ggggaaatac tatcaactga accaaaaata 1980

acttgtcata ggcagagata aagccagaaa cactctacac atgccagatg acatctggag 2040

aaaagggtgc taagggaagc gtttggcagc aagatatgat tgtaagggtg tgtcccttga 2100

gttcaatgtc tgcctatttc tgatgggtct aaagcaacat ggagttactg tgcagcagaa 2160

ctctcagtaa agacaccatt tgccttggca atcctcaaaa agcttcaata gcagattgct 2220

tcagaccatc tgtagtccgt ccttttctca tctggatgtt gtttggcttc tgtgcgaaag 2280

attggtggag tgtccagta gatatcatgg tgggtgtgtga tcagagtccc aaggaacctg 2340

aatgagccaa ggtgcccagc atgaagtcaa acaaaagcct tgacatgagt ttgccatgaa 2400

atagcgaaga gagagtggaa gagaggagcc aatcactgtg gggcagtgcc accctgaggg 2460

cacttagggt atgggggttg tgcttaaata catcacagat ccaggtactg aatgggagga 2520

agtgtgggtg atttccaatc tcattgacct tatgttcagg gacttgaacg gaagatgttt 2580

cttgtgttgc ctaagtggta ttcagtctac cagactctgc aacttgcac ttcaaactct 2640

tggtaaagag atgtggatgg tgtcagagaa ggcaaaggcc tgcagtggat tgaagaggct 2700

tgcaagcagt tctgtttcta ggatgtgggc ttcacacagaa gacactcggc caccacttag 2760

ctagtctaaa cctcagggtt cctcagccca tcatacccca acttgaggga ctgacatcaa 2820

ggagtagact ggagaaacag ccctcccatc aagtaacctc ttgttctctc ctgctccatc 2880

tgcaactatag aagtgttaata attagacata cttggcaaaa tggctaattg atttggtaac 2940

agaagcatga gccataacaa tggaagatct agttatcatg actgaacagc ttaacattca 3000

attcccttct ctaagagaag ctgtgaaatc ctacatatta tttaaagtta accaaatcaa 3060

tgtaaagggg gtaggagac agtgtgtacc tatgcacgta ttttatgtt ttgcttgtgt 3120
tccagtctcg gtcatttggt tccattttca agcaatttat ttgaagagcc attgcactag 3180
cctgatgtat actgcaatga gcttctttga taaaatgaaa cttaaatttt tctcgaccat 3240
ttcaccgtgc ctctacttc attttttgcc agaaaatctc acatccaaca aaacaaaaca 3300
aaaaccctga attagtgggc ttgaaaagg aaaaagcagg gctttgaaaa agtagatcac 3360
acatcagtta agactcctgc ttctctatta gtcagggtgt cttggattca gtctggagta 3420
ggcagagctt aagggttttt aagtcctgac ccaaagaaat gatctagcct gaaagttag 3480
agcaaaggac taatgtttac ttttaaagga atttcttgat ttttttaaaa aacttcatta 3540
aagttaaata cccaatgga caaatcata atcttgtaa tcgttattac taaacttttt 3600
aaaaaatgtc ccaatttaca attaaataaa ttactttctc agtatattct ggtctggtca 3660
tggtattgtc atttcctccc aaagatattc aaaattgtca attagagaat tttaggtttt 3720
cagactcaga aaagtcctca cgccc 3745

<210> 35

<211> 244

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)..(60)

<223> 5' flanking sequence

<220>

<221> misc_feature

<222> (61)..(178)

<223> human untranslated exon 1

<220>

<221> Intron

<222> (179)..(244)

<400> 35

agcccggccg gccggccac gggcgggagg acgcgcctcc gctcgggagg aggcggcgcg 60

gtggctgata agagcgcgta gggcttcgcc ggggccggga gctgggagcg gtcctgctca 120

gccagctca ccgcgcgccg gccctcggcg ccctcggcgc cctgggtctg cggatcaggt 180

gggtcccgcg gggagccgcc caggtcccg gaggccacga gcaggacacg gacggggggc 240

nnnn 244

<210> 36

<211> 217

<212> DNA

<213> Homo sapiens

<220>

<221> Intron

<222> (1)..(60)

<220>

<221> misc_feature

<222> (61)..(149)

<223> human untranslated exon 4

<220>

<221> Intron

<222> (150)..(217)

<400> 36

ctcttgaagt tcattgattt aatctgttct ctttttttct cccctcttct tttttcctag 60

gagaaaataa tgaatgtcaa aggaaaagta attctgtcaa tgctggttgt ctcaactgtg 120

atcattgtgt ttgggaatt tatcaacagg taattatgaa acatgatgaa gtgatgtgga 180

tgaaaatact gctttgattc taccctacta gtatnnn 217

<210> 37

<211> 165

<212> DNA

<213> Homo sapiens

<220>

<221> Intron

<222> (1)..(60)

<220>

<221> misc_feature

<222> (61)..(96)

<223> human untranslated exon 5

<220>

<221> Intron

<222> (97)..(165)

<400> 37

aatcgccctt ctcagaatta aaagtaacat gatatgtttt tatttctttt ttgcttttag 60

cacagaaggc tctttcttgt ggatatatca ctcaaagtgc ttgaattct agatttctag 120

gggatgtttc ccacagccac tctggcaccc cctacagtcc annnn 165

76

<210> 38
<211> 193
<212> DNA
<213> Homo sapiens

<220>
<221> Intron
<222> (1)..(60)

<220>
<221> misc_feature
<222> (61)..(126)
<223> human untranslated exon 6

<220>
<221> Intron
<222> (127)..(193)

<400> 38
accctaagtt tggggacacc acattttcta aaaatatttg taaacttttt catttcttag 60

aaacccagaa gttgatgaca gcagtgtca gaagggtgg tggtttctga gctggtttaa 120

caatgggtaa ggcgatcag acagcagtcg gtgtttgcc accgcctgg tgcttcaga 180

gggtccnnnn nnn 193

<210> 39
<211> 242
<212> DNA
<213> Homo sapiens

<220>
<221> Intron
<222> (1)..(60)

77

<220>

<221> misc_feature

<222> (61)..(176)

<223> human untranslated exon 7

<220>

<221> Intron

<222> (177)..(242)

<400> 39

tctttgacca ccgcaatcac ctccctgcc ttacctggtt tactttccct ttgtacttag 60

gatccacaat tatcaacaag gggaagaaga catagacaaa gaaaaaggaa gagaggagac 120

caaaggaagg aaaatgacac aacagagctt cggtatggg actggtttaa tccaaagtaa 180

gaaaagcggc gtcactccct gtgcagcaaa tccatggccc tgcagggggt ggtgtggcnn 240

nn

242

<210> 40

<211> 487

<212> DNA

<213> Homo sapiens

<220>

<221> Intron

<222> (1)..(60)

<220>

<221> misc_feature

<222> (61)..(487)

<223> a version of human untranslated exon 8h

<400> 40

78

atagaatatt ttaattttta attcaacata aatttttaag ggtgctgttt tttcttccag 60
cttgaaggaa tccgaataac taaactggac tctggttttc tgactcagtc cttctagaag 120
acctggactg agagatcatg cggttaagga gtgtgtaaca ggcggaccac ctgttgggac 180
tgcgagattc tcaaggggaa ggactgggtc tcattttctcc catctcagcg cttagcagga 240
tgacctggta tagagcaggg aactgggaaa tgtgggtcag gggatcagac actccagttg 300
ggtctttttat ataaattaaa tggcaaaagg ctccataccc ttctccttct ttcctaccct 360
ccactttatc tgcaaaatgg gaatgatgat aacaccctact tcatagaatg gtcatgaaga 420
tcaaatagaga gaataaaagt caagcactta gcctctggtg cacaataagt attaaataag 480
tataacct 487

<210> 41

<211> 454

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)..(380)

<223> a version of the human untranslated exon 8h

<220>

<221> Intron

<222> (381)..(454)

<400> 41

attcctcctt ttcctttttt aaaaataata ttaccaaagc tccagcttat acacatttac 60
aagacttagc tagtgggcta tgtagagct actaaaagat ctttgacaag ctaaaactaa 120

gatgcaatga atgagggtgta acgaacaaga gaggttttaag ttcagaaatg gttacagaag 180
tataagacag ctgtgtgggt gttttttggt ttttggtttc tggtttaca tctcgtcatt 240
caacaaagat gggagtttta tagaactaaa agcaccatgt aagctactaa aaacaacaac 300
aaaaaaggct catcatttct cagtctgaat tgacaaaaat gccaatgcaa ataaaaatga 360
ttacttttta tttttcaacg ttgtttgttt atttatttat ttcgagatgg agtttcactc 420
ttgttgccct ggctggagtg cagtggcgcn nnnn 454

<210> 42

<211> 2848

<212> DNA

<213> Homo sapiens

<220>

<221> Intron

<222> (1)..(65)

<220>

<221> misc_feature

<222> (66)..(2676)

<223> human untranslated exon 9

<220>

<221> misc_feature

<222> (2677)..(2848)

<223> an inter-gene sequence

<400> 42

ttcagcttgt ggtttctttc aggaatccca gaggataaat gttttgcttt tcttctttgt 60

ttcagatata atgatcatta cttggaggag ttcataacat ctgctaatag gtacttcatg 120
gttggccaca aagtcattt ttacatcatg gtggatgatg tctccaagct gccgtttata 180
gagctgggtc ctctgcattc cttcaaaatg tttgaggcca agccagagaa gaggtggcaa 240
gacatcagca tgatgcgtat gaagatcact ggggagcaca tcttgccca catccaacac 300
gaggtcgact tctctttctg catggatgtg gaccaggtct tccaagacca ttttggggtg 360
gagaccctag gccagtcagt ggctcagcta caggctggcg gtacaaggca gatccctatg 420
actttaccta ggagaggtgg aaagagtcag caggatacat tccatttggc caggggattt 480
ttattaccat gcagccattt ctggaggaac acccattcag gttctcaaca tcaccagga 540
gtgctttaag ggaatcctcc tggacaagaa aaatgacata gaagccaagt ggcattgatga 600
aagccaccta aacaagtatt tctttctcaa taaacctct aaaatcttat cctaaaaata 660
ctgctgggat tatcatatag gcctgccttc agatattaaa actgtcaagt gatcgtggca 720
gacaaaagag tataatttgg ttagaaataa tgtctgactt caaattgtgc cagtagattt 780
ctgaatttaa gagagagaat attctggcta cttcctcaga aaagtaacac ttaattttaa 840
cttcaaaaaa tactaatgaa acaccaacag ggcaaaaaca taccattcct ccttgtaact 900
tggggctttg taatgtggaa gaatgaatct agggcaatca gatataaatt ccagtgatt 960
tcttatctat tctgggtttg ggggaaatac tatcaactga accaaaaata acttgtcata 1020
ggcagagata aagccagaaa cactctacac atgccagatg acatctggag aaaagggtgc 1080
taagggaagc gtttggcagc aagatatgat tgtaagggtg tgtcccttga gttcaatgtc 1140
tgcctatttc tgatgggtct aaagcaacat ggagttactg tgcagcagaa ctctcagtaa 1200

agacaccatt tgccttggca atcctcaaaa agcttcaata gcagattgct tcagaccatc 1260

tgtagtccgt ccttttctca tctggatgtt gtttggttc tgtgcgaaag attggtggag 1320

tgccccagta gatatcatgg tgggtgtgta tcagagtccc aaggaaacctg aatgagccaa 1380

ggtgcccagc atgaagtcaa aacaaagcct tgacatgagt ttgccatgaa atagcgaaga 1440

gagagtggaa gagaggagcc aatcactgtg gggcagtgcc accctgaggg cacttaggg 1500

atggggttgg tgcttaaata catcacagat ccagggtactg aatgggagga agtgtgggtg 1560

atttccaatc tcattgacct tatgttcagg gacttgaacg gaagatgttt cttgtgtgc 1620

ctaagtggta ttcagtctac cagactctgc aacttgcac ttcaaatcct tggtaaagag 1680

atgtggatgg tgcagagaa ggcaaaggcc tgcagtggat tgaagaggct tgcaagcagt 1740

tctgtttcta ggatgtgggc ttcacagaa gacactcggc caccacttag ctagtctaaa 1800

cctcagggtt cctcagccca tcatacccca acttgaggga ctgacatcaa ggagtagact 1860

ggagaaacag ccctcccatc aagtaacctc ttgttctctc ctgctccatc tgcactatag 1920

aagtgttaata attagacata cttggcaaaa tggctaattg atttggtaac agaagcatga 1980

gccataacaa tggaagatct agttatcatg actgaacagc ttaacattca attcccttct 2040

ctaagagaag ctgtgaaatc ctacatatta tttaaagtta accaaatcaa tgtaaaggga 2100

gttaggagac agtgtgtacc tatgcacgta tatttatgtt ttgcttgtgt tccagtctcg 2160

gtcatttgtt tccattttca agcaatttat ttgaagagcc attgcactag cctgatgtat 2220

actgcaatga gcttctttga taaaatgaaa cttaaatttt tctcgaccat ttcaccgtgc 2280

ctcctacttc attttttgcc agaaaatctc acatccaaca aaacaaaaca aaaaccctga 2340

attagtgggc tttgaaaagg aaaaagcagg gctttgaaa agtagatcac acatcagtta 2400
agactcctgc ttctctatta gtcaggttgt cttggattca gtctggagta ggcagagctt 2460
aagggttttt aagtctgac ccaaagaaat gatctagcct gaaagttag agcaaaggac 2520
taatgtttac ttttaaagga atttcttgat ttttttaaaa aacttcatta aagtttaaata 2580
ccccaatgga caaatcata atcttgtaa tcgttattac taaacttttt aaaaaatgtc 2640
ccaatttaca attaaataaa ttactttctc agtatattct ggtctggtca tggattgtgc 2700
atttcctccc aaagatatc aaaattgtca attagagaat ttaggtttt cagactcaga 2760
aaagtccctca cgcccttctg aaaatgtgtc cactattaca gaaatagaac agacttgga 2820
ttcccaaatt tttgtttgtt tttnnnnn 2848

<210> 43
<211> 2303
<212> DNA
<213> Rhesus monkey

<220>
<221> misc_feature
<222> (1)..(44)
<223> This is exon 1

<220>
<221> misc_feature
<222> (45)..(159)
<223> This is exon 2

<220>
<221> misc_feature

<222> (160)..(278)
<223> This is exon 3

<220>
<221> misc_feature
<222> (279)..(367)
<223> This is exon 4

<220>
<221> misc_feature
<222> (368)..(403)
<223> This is exon 5

<220>
<221> misc_feature
<222> (404)..(469)
<223> This is exon 6

<220>
<221> misc_feature
<222> (470)..(584)
<223> This is exon 7

<220>
<221> misc_feature
<222> (585)..(2260)
<223> This is exon 9

<220>
<221> misc_feature
<222> (543)..(545)
<223> This is an early stop codon

<220>

<221> misc_feature

<222> (1276)..(1278)

<223> This is a stop codon

<220>

<221> misc_feature

<222> (1225)..(1260)

<223> This is a polyadenylation signal

<400> 43

gctcgtcgcg cgccggctct ggggtgccagg gttctgcgga tcaggagttg aaaccagcat	60
cttcccttca tctgagtcct gcctccttct gcagaaggga gctcaaaaga actttgttgt	120
tttgctttt actctggggt gaaagcaaca gacgataagg atctcactct gtcgccaag	180
ctggagtgca gtggcttgat tacagctcac tgtagcctgg accttccaag gctctgggtg	240
atcttcctac ctgagcttcc ccagtagctg gactacagga gaaaataatg aatgtcaaag	300
gaaaagtaat tctgtcaatg ctggttgtct caactgtgat cattgtgttt tgggaatata	360
tcaatagccc agaaggttct ttcttgggga tgtatcgctc aaaaaacca gaggttgatg	420
acagcagtgc tcagaagagc tgggtggttc cgagctggtt taacaatggg atccacaatc	480
atcaacaaga ggaagaagac atagacaaaa aagaggaaga gaggagacca aagaaaggaa	540
gatgacacaa cagagcttcg gctatgggac tgatttaatc caaaatatat tgagcattac	600
ttggaagagt tcataacacc tgctaatagg tacttcaagg tcggccacaa agtcatattt	660
tacattatag tggatgatgt ctccaagggtg ctgtttatag agctgggtcc tctgcattcc	720

ttaaaagtgt ttgaggtcaa gccagagaag aggtggcaac acatcagcat gatgcctgtg 780
aagatcatca gggagcacat cttggccac atccaacacg aggtcgactt cctcttctgc 840
atggatgtag accaggtctt ccaagacaat tttggggtga agaccctagg tcagtcagtg 900
gctcagctac agccctggtg gtacaaggca gatcctgatg actttaccta ggagaggcag 960
aaagagtcag cagcatgcat tccatttggc caggaggatt tttattacca cacagccatt 1020
tttgaggaa caccattca ggttctcaac atccccagg agtgctttaa gagaatcctc 1080
ctggaaaaga aaaatgacat agaagctgag tggcatgatg aaagccacct aaaccagtat 1140
ttccttctca acaaaccctc taaaatctta tcctagaat actgctggga ttatcatatc 1200
agcctgcctt cagatattaa aactgtcaag cggctgtggc agacaaaaga gtataatttg 1260
gttagaaata tcatctgact tcaaattgtg ccagtagatt tctgaatttg agagaggagt 1320
attctggctg cttcctcaga aaagtaacac ttaattttta gttaaaaaaa atactaatga 1380
aacaccaaca tggcaaacac ataccattcc ttcttgtaac ttgaggcttt gtaatgtggg 1440
agaatgaatc tagggtaatc agatgtaat tcccagtgat ttcttatcta ttttgggttt 1500
gggggaaata ctatcaactg aacaaaaaag aacttgtcat aggcaaagat aaagccagaa 1560
acactctaca catgccacat aacatctgga gaaaagggtg ctaagggaag cgtttggcag 1620
caagatatga ttgtaagggg ttgtcccttg agttcaatgc ctgcctatct ccaatggatc 1680
taaaacaacg tgaagtact gtgcagcaga gctctcagta aggacacat ttgccttggc 1740
aatcctcaaa attcttcaat agcagattgt ttcaggccat ctgtagtctg tccttttctc 1800
atcaggatgt tgtttggctt ctgtgcgaaa aattgggtga gtgtcctggt agatattgaa 1860

actaggcctc atatagaaaa aattaacacc aggtggctct ggatagagtc ccgccctgcc 1920
tcgatgagga cccaccctga taggggccca ccctgccaat tccgagaaac aacctcatgg 1980
ggccccaccc tgccaattcc ggggggccca ccctgcctcg aagttcccgg aatcaacaac 2040
tccaggaaaa aacctcataa ggtcctgctc taaccaatta gcataagacg ccttgctcag 2100
gccatagcta gacccaatca ttttgcgcct taagctttgt ttgaatttcg cgccctaagc 2160
tgtgtttgaa cttgtgtttg cctatataaa cagcctgtaa caagcagtcg ggggccagg 2220
gccaacttag agcttgggac cctagcgcg tagtaataaa taactctctg ctgcgaaaaa 2280
aaaaaaaaa aaaaaaaaaa aaa 2303

<210> 44

<211> 2630

<212> DNA

<213> Rhesus monkey

<220>

<221> misc_feature

<222> (1)..(44)

<223> This is exon 1

<220>

<221> misc_feature

<222> (45)..(159)

<223> This is exon 2

<220>

<221> misc_feature

<222> (160)..(278)

<223> This is exon 3

<220>

<221> misc_feature

<222> (279)..(367)

<223> This is exon 4

<220>

<221> misc_feature

<222> (368)..(403)

<223> This is exon 5

<220>

<221> misc_feature

<222> (404)..(469)

<223> This is exon 6

<220>

<221> misc_feature

<222> (470)..(584)

<223> This is exon 7

<220>

<221> misc_feature

<222> (585)..(911)

<223> This is exon 8

<220>

<221> misc_feature

<222> (912)..(2587)

<223> This is exon 9

<220>

<221> misc_feature

<222> (288)..(290)

<223> This is a putative start codon

<220>

<221> misc_feature

<222> (543)..(545)

<223> This is a putative stop codon

<220>

<221> misc_feature

<222> (1603)..(1605)

<223> This is a putative stop codon

<220>

<221> misc_feature

<222> (2582)..(2587)

<223> polyadenylation signal

<400> 44

gctcgctgcg cgccggtcct gggtgccagg gttctgcgga tcaggagttg aaaccagcat 60

cttccttca tctgagtcct gcctccttct gcagaaggga gtcaaaaga actttgttgt 120

tttgcccttt actctggggt gaaagcaaca gacgataagg atctcactct gtcgcccaag 180

ctggagtgca gtggcttgat tacagctcac ttagcctgg accttccaag gctctgggtg 240

atcttctac ctcagcttcc ccagtagctg gactacagga gaaaataatg aatgtcaaag 300

gaaaagtaat tctgtcaatg ctggttgtct caactgtgat cattgtgttt tgggaatata 360

tcaatagccc agaaggttct ttcttgggga tgtatcgctc aaaaaacca gaggttgatg 420

acagcagtgc tcagaagagc tggtggtttc cgagctggtt taacaatggg atccacaatc 480

atcaacaaga ggaagaagac atagacaaaa aagaggaaga gaggagacca aagaaaggaa 540

gatgacacaa cagagcttcg gctatgggac tgatttaatc caaagaaacg cccagaggtg 600

gtgagagtga ccagatggaa ggcaccggtt gtgtggaaag gcacttacia caaagccatc 660

ctaggaaatt attatgccaa acagaaaatt acggtgggat tgaaggcttt tgctattgga 720

agtgggtgtc actgatgaaa ctgtccttga ctatttcttg ttccactgtc aagacatttt 780

tgtggagact cctgaactga tggaggccag ccatgatttt ttgatttatt agatagaaga 840

atgttttcat ggaactgttt tagtctcctt tctgctgagg ccctaaaatg ctgagaacaa 900

aataagagta gatatatgga gcattacttg gaagagttca taacacctgc taataggtac 960

ttcaaggctg gccacaaagt catattttac attatagtgg atgatgtctc caagggtgctg 1020

tttatagagc tgggtcctct gcattcctta aaagtgtttg aggtcaagcc agagaagagg 1080

tggcaacaca tcagcatgat gcctgtgaag atcatcaggg agcacatctt ggcccacatc 1140

caacacgagg tcgacttctt cttctgcatg gatgtagacc aggtcttcca agacaatttt 1200

ggggtgaaga ccctaggtca gtcagtggct cagctacagc cctggtggta caaggcagat 1260

cctgatgact ttacctagga gaggcagaaa gagtcagcag catgcattcc atttgccag 1320

gaggattttt attaccacac agccattttt ggaggaacac ccattcaggt tctcaacatc 1380

ccccaggagt gctttaagag aatcctcctg gaaaagaaaa atgacataga agctgagtgg 1440

catgatgaaa gccacctaaa ccagtatttc cttctcaaca aaccctctaa aatcttatcc 1500

ctagaataact gctgggatta tcatatcagc ctgccttcag atattaaaac tgtcaagcgg 1560
tcgtggcaga caaaagagta taatttggtt agaaatatca tctgacttca aattgtgcc 1620
gtagatttct gaatttgaga gaggagtatt ctggctgctt cctcagaaaa gtaacactta 1680
attttaagtt aaaaaaata ctaatgaaac accaacaagg caaacacata ccattccttc 1740
ttgtaacttg aggctttgta atgtgggaga atgaatctag ggtaatcaga tgtaaattcc 1800
cagtgatttc ttatctatct tgggtttggg ggaaatacta tcaactgaac caaaaagaac 1860
ttgtcatagg caaagataaa gccagaaaca ctctacacat gccacataac atctggagaa 1920
aagggtgcta agggaagcgt ttggcagcaa gatatgattg taaggggttg tcccttgagt 1980
tcaatgcctg cctatttcca atggatctaa aaacaactga agttactgtg cagcagagct 2040
ctcagtaagg acaccatttg ccttggaat cctcaaaatt cttcaatagc agattgtttc 2100
aggccatctg tagtctgtcc tttctctatc aggatgttgt ttggcttctg tgcgaaaaat 2160
tggtggagtg tcctggtaga tattgaaact aggcctcata tagaaaaaat taacaccagg 2220
tggtcttga tagagtcccg ccctgcctcg atgaggaccc accctgatag ggtcccaccc 2280
tgccaattcc gagaaacaac ctcatggggc cccaccctgc caattccggg ggtcccaccc 2340
tgctcgaag ttcccgaat caacaactcc aggaaaaaac ctcataaggt cctgctctaa 2400
ccaattagca taagacgcct tgctcaggcc atagctagac ccaatcattt tgcgccttaa 2460
gctttgtttg aatttcgcgc cctaagctgt gtttgaactt gtgtttgcct atataaacag 2520
cctgtaacaa gcagtcgggg tcccagggcc aacttagagc ttgggaccct agcgcgctag 2580
taataaataa ctctctgctg cgaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2630

<210> 45
<211> 35
<212> DNA
<213> Artificial/Unknown

<220>
<221> misc_feature
<222> ()..()
<223> Antisense primer for cloning porcine exon 4

<400> 45
ctgttgatgt attcccaaaa cacaaccatt acagt

35

<210> 46
<211> 27
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> ()..()
<223> Antisense primer for cloning porcine exon 4

<400> 46
agacaagcag cattgacaga accactc

27

<210> 47
<211> 25
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature

<222> ()..()

<223> Antisense primer for cloning porcine exon 2

<400> 47

ctcatcctct gcttctctcc cccca

25

<210> 48

<211> 26

<212> DNA

<213> artificial sequence

<400> 48

ccccccagag taaaaggcga aacaag

26

<210> 49

<211> 25

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 2

<400> 49

aacgcagcac cttcccttcc tccca

25

<210> 50

<211> 25

<212> DNA

<213> artificial sequence

<400> 50

93

cttgtttcgc cttttactct ggggg

25

<210> 51

<211> 23

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 1

<400> 51

gccactgttc cctcagccga gga

23

<210> 52

<211> 24

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 1

<400> 52

cgagcgcacc cagcttctgc cgat

24

<210> 53

<211> 24

<212> DNA

<213> artificial sequence

94

<220>

<221> misc_feature

<222> ()..()

<223> Antisense primer for cloning porcine exon 1

<400> 53

tgcgctcggg gatggccctc tcct

24

<210> 54

<211> 24

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Antisense primer for cloning porcine exon 1

<400> 54

ggcgctcctcg gctgagggaa cagt

24

<210> 55

<211> 28

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 1A

<400> 55

cagaacaact tctgaagcct aaaggatg

28

<210> 56

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 1

<400> 56

caaatggtgg atcggacctc ccaggct

27

<210> 57

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 1

<400> 57

agtactgggt gatagacccc actccac

27

<210> 58

<211> 25

<212> DNA

<213> artificial sequence

<220>

96

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 1

<400> 58

gcgcagggt cggggcccc tccct

25

<210> 59

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 9

<400> 59

ctgggattat catataggca tgtctgt

27

<210> 60

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> Sense primer for cloning porcine exon 9

<400> 60

agagtattac tctggctact tctccag

27

<210> 61
<211> 27
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> ()..()
<223> primer for identifying 5' flanking region of murine exon 1

<400> 61
ctgagagcgc gaggtcttca gcagaat

27

<210> 62
<211> 28
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> ()..()
<223> primer for identifying 5' flanking region of murine exon 1

<400> 62
cttctcattc caagaagagt cttacaag

28

<210> 63
<211> 27
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature

<222> ()..()

<223> primer for identifying 3' flanking region of murine exon 1

<400> 63

cctgcctttt cttagctggc tgacacc

27

<210> 64

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying 3' flanking region of murine exon 1

<400> 64

cttgtagact cttcttgga tgagaag

27

<210> 65

<211> 27

<212> DNA

<213> artificial sequence.

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying 5' flanking region of murine exon 2

<400> 65

catcgtcagc tgtgttcct ccaaagc

27

99

<210> 66

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying 5' flanking region of murine exon 2

<400> 66

aaagcaaccg agcttctgtc gagctct

27

<210> 67

<211> 38

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exons 2 and 3

<400> 67

gtaccttcttctcctctgtc gagccctgcc tccttcgg

38

<210> 68

<211> 35

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

100

<223> primer for identifying murine exons 2 and 3

<400> 68

agatcttgag gatccaagac ttgtttctga cttgg

35

<210> 69

<211> 34

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()...()

<223> primer for identifying murine exons 3 and 4

<400> 69

gctgactttg aactcaagag atctgcttta cccc

34

<210> 70

<211> 28

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()...()

<223> primer for identifying murine exons 3 and 4

<400> 70

ctgttgacat attccccaaa caccgacaa

28

<210> 71

101

<211> 30
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> ()..()
<223> primer for identifying murine exons 4 and 5

<400> 71
gtcaagggaa aagtaatcct gttgatgctg

30

<210> 72
<211> 27
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> ()..()
<223> primer for identifying murine exons 4 and 5

<400> 72
tatccacaag aaagagccgt ctgggct

27

<210> 73
<211> 27
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> ()..()
<223> primer for identifying murine exons 5 and 6

<400> 73

agcccagacg gctctttctt gtggata

27

<210> 74

<211> 34

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exons 5 and 6

<400> 74

ccagcttggg aaccaccagt ccttctgccca tctg

34

<210> 75

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exons 6 and 7

<400> 75

ttccagaggt tggtgagaac agatggc

27

<210> 76

<211> 33

103

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exons 6 and 7

<400> 76

gcgatctcca tttctaccct tttctctccg tcc

33

<210> 77

<211> 28

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exon 7

<400> 77

caagaagaca acgtagaagg acggagag

28

<210> 78

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exon 7

<400> 78

tcgcattgaa gagcctcagc tatggga

27

<210> 79

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying exon 8

<400> 79

ccacagtgag tttctgtgtg gcgatgt

27

<210> 80

<211> 28

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exon 8

<400> 80

agagctgtgt cataagtgcc ttcccaca

28

<210> 81

<211> 27

<212> DNA

105

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exon 8

<400> 81

gatgttttga cagtgacccc gtggaag

27

<210> 82

<211> 28

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exon 8

<400> 82

tgtgggaagg cacttatgac acagctct

28

<210> 83

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exon 9

106

<400> 83

agagggttca ggtgcacgac aggcac

27

<210> 84

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying murine exon 8

<400> 84

gtacatgtca gcagactcca gaaagtc

27

<210> 85

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying 3' flanking region of murine exon 9

<400> 85

gactttctgg agtctgctga catgtac

27

<210> 86

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying 3' flanking region of murine exon 9

<400> 86

gatgcctgtc gtgcacctga accctct

27

<210> 87

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying 3' flanking region of murine exon 9

<400> 87

aggccattgc accatcttgg tgaacag

27

<210> 88

<211> 28

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for identifying 3' flanking region of murine exon 9

<400> 88

108

gatcttacct ttgtccacag ggctctac

28

<210> 89

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for obtaining murine promoter

<400> 89

ccaatgcac ttttcccagt gggctct

27

<210> 90

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for isolation of transcription initiation site

<400> 90

cccagaacag atctgactgc ctctttc

27

<210> 91

<211> 27

<212> DNA

<213> artificial sequence

109

<220>

<221> misc_feature

<222> ()..()

<223> primer for isolation of transcription initiation site

<400> 91

agttttgctt gtctgggcca ctatcgg

27

<210> 92

<211> 27

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for isolation of transcription initiation site

<400> 92

gactggagag agtgctgtcc tccttgc

27

<210> 93

<211> 29

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for cloning Rhesus alpha 1,3 GT

<400> 93

gagggtcaagc cagagaagag gtggcaaca

29

<210> 94

<211> 30

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for cloning Rhesus alpha 1,3 GT

<400> 94

gacttcctct tctgcatgga ttagaccag

30

<210> 95

<211> 29

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<222> ()..()

<223> primer for cloning Rhesus alpha 1,3 GT

<400> 95

atgtcgagaa cctgaatggg tgttctctcc

29

<210> 96

<211> 30

<212> DNA

<213> artificial sequence

<220>

111

<221> misc_feature

<222> ()..()

<223> primer for cloning Rhesus alpha 1,3 GT

<400> 96

ctggccaaat ggaatgcatg ctgctgactc

30